Journal of
Clinical Psychology

OCTOBER 1950
Volume VI Number 4

Contents

Intrafamilial Personality Configurations Among Children with Primary Behavior Disorders and their Parents: A Rorschach Investigation. Woodrow W. Morris and Alma L. Nichols 309
Objective Studies of Children's Drawings of Human Figures. I. Sex Awareness and Socio-Economic Level. Arthur Weider and Paul A. Noller, University of Louisville School of Medicine 315
Personality Characteristics in Hypnotically-Induced Hostility. Robert M. Counts and Ivan N. Mensh, Washington University School of Medicine 325
A Study of the Effect of Color on Productivity and Reaction Time in the Rorschach Test. Raphael J. Dubrovner, William J. Von Lackum and Hudson Jost, University of Tennessee 331
A Study of The Diagnostic Possibilities of the CVS with a Group of Diagnostic cases. Joseph D. Matatazzo, Northwestern University 337
Validation Studies of the Rorschach Test: Shyness and Gregariousness in the Normal Supreior Adult. Wayne H. Holtzman, University of Texas 343
Validation Studies of the Rorschach Test: Impulsiveness in the Normal Superior Adult. Wayne H. Holtzman, University of Texas 348
Rorschach Content in Hypnosis: Chronological Age Regression. Margaret Mercer and R. W. Gibson, Veterans Administration, Coatesville, Pa. 352
A Study of Wechsler's Concept of "Normal Deterioration" in Older Mental Defectives. Gerard J. Bensberg and William Sloan, Lincoln State School and Colony 359
Judgment of Adjustment from Drawings: The Applicability of Rating Scale Methods. George W. Albee and Roy M. Hamlin, University of Pittsburgh 363
The Hanfmann-Kasanin Concept Formation Test as a Measure of Rigidity in Relation to College Aptitude and Achievement. Ralph A Baker and Donald G. Doehring, University of New Mexico 365
Group Discussion as a Method of Patient Orientation in a Neuropsychiatric Hospital. James O. Palmer, Washington University School of Medicine 369
The Negro Patient in Psychotherapy. Ralph W. Heine, Veterans Administration, Chicago 373
An Attempted Experimental Validation of the Thematic Apperception Test. James S. Calvin, University of Kentucky, and Leo C. Ward, Veterans Administration, Atlanta 377
Szondi Results in Group Testing with College Students. David Cole and Eugene Roberts, Occidental College 381

(Continued on next page)

JOURNAL OF CLINICAL PSYCHOLOGY

EDITORIAL BOARD

FREDERICK C. THORNE, *Editor*
Medical College, University of Vermont

ELEANOR C. THORNE, *Business Manager*

GENERAL INFORMATION

The *Journal of Clinical Psychology* is an independent journal dedicated to the advancement of the clinical method in psychology. Although primarily a scientifically oriented professional journal limited to publication of original research reports and authoritative theoretical articles, it aims to foster the promotion and expansion of clinical psychology as an applied science. Special sections in each issue will be devoted to news, comments, and letters to the editors on matters pertinent to the profession. Communications and articles submitted for publication are welcome and should be addressed to the Editorial Office, 5 Pearl Street, Brandon, Vermont.

REPRINTS are furnished on order only and must be requested on the order blank provided when galley proofs are returned. Change of address notices should be sent one month before moving and should include old and new addresses. Undelivered copies will be remailed only at subscribers' expense.

The *Journal of Clinical Psychology* is published quarterly, appearing in January, April, July and October, at Burlington, Vermont, by Frederick C. Thorne, M.D. Subscription price $6.50; Canadian subscriptions $7.00; foreign subscriptions $7.50 including postage; single copies $2.00. Entered as second class matter January 15, 1945, at the post office at Burlington, Vermont, under the Act of March 3, 1879. Additional entry at the post office at Brandon, Vermont and Rutland, Vermont. *Title registered, U. S. Patent Office.* Editorial and Business Offices, 5 Pearl Street, Brandon, Vermont.

Copyright, 1950, Frederick C. Thorne, M.D.

Contents

(*Concluded from Front Cover*)

The Test Situation and the Problem of Prediction. Seymour B. Sarason, Yale University 387
Mental Age Scores for the Wechsler Intelligence Scale for Children. William B. Grove, Phoenix Elementary Schools 393
The Wechsler-Bellevue Test and the Diagnosis of Psychopathic Personality. Milton S. Gurvitz, Hillside Hospital, Bellerose, N. Y. 397
A Study of the Differential Responses on the Vocabulary Sub-test of the Wechsler Intelligence Scale for Children. Chalmers L. Stacey and Bernard Portnoy, Syracuse University 401
Report on a Method of Predicting Success in Psychotherapy. Theodore H. Blau, Pennsylvania State College 403
Differential Scoring of the Rosenzweig Picture-Frustration Study. George Spache, Chappaqua, N. Y. 406
A Comparison of Two Methods of Administering the Digit Span Test. Richard L. Newton, University of Pittsburgh 409
A Note Regarding the Suggested Use of the Bender Visual Motor Gestalt Test as a Measure of School Readiness. Marcella Vig Baldwin, Bureau for Colored Children, Philadelphia, Pa. 412
A Note on the Presentation of Personality Inventory Scores. Kenneth H. Ives, New School for Social Research 415
A Note on T. A. T. Administration. Mary Alice White and Hanna Schreiber, New York Hospital, Westchester Division 417

Editorial Opinion 418
Book Reviews 420
Subject Index 423
Author Index 426

JOURNAL OF
CLINICAL PSYCHOLOGY

VOL. VI OCTOBER 1950 NO. 4

INTRAFAMILIAL PERSONALITY CONFIGURATIONS AMONG CHILDREN WITH PRIMARY BEHAVIOR DISORDERS AND THEIR PARENTS: A RORSCHACH INVESTIGATION[1]

WOODROW WILBERT MORRIS AND ALMA L. NICHOLAS[2]

The Iowa State Psychopathic Hospital and the State University of Iowa College of Medicine

INTRODUCTION

While much has been reported with regard to the more outstanding characteristics of psychopathic personality and of children suffering primary behavior disorders, at least two important problems have been seriously neglected. One of these is concerned with the systematic evaluation of the personality structure and dynamics of such patients and the other with the genesis of such a personality structure. It will be the purpose of the present communication to study these two aspects of the problem by means of the Rorschach Method of psychodiagnosis.

The two groups of patients are considered together here since in previously reported studies in this series they have proved to be clinically similar. Thus, Gottlieb, *et al* state:

> These two diagnostic groups were separated arbitrarily on the basis of age, disturbances in patients of 16 years or above being considered as psychopathic personality and disturbances in patients of 15 years or below as primary behavior disorders. Thus, on a descriptive level, the concept of psychopathic personality became an extension of the concept of primary behavior disorders in children . . . . [2].

Studies of similar nosological groups using various psychological tests have centered primarily toward an attempt to discover "signs" which might be of assistance in diagnosis rather than any attempt to understand better the structure of the personality thus designated. Thus Heuser[5] reports a Rorschach investigation of 28 military cases diagnosed as psychopathic personality and found 13 "signs" which he felt were of value in the Rorschach picture of the psychopath. Lindner[8] on the other hand reports almost complete failure in his attempt to delineate "signs" of psychopathy with the exception of "sequence analysis" which he highly recommends.

The function of the clinical psychologist in his use of such promising techniques as the Rorschach Method, however, is far more important than the demonstration of "signs" which will discriminate pathological personalities from normal ones. While the latter function has its values, it too frequently leads to the misapprehension that Rorschach analyses can be done by the so-called "sign approach" leading merely to a diagnosis. It is obvious to those who have made extensive clinical use of the Rorschach that its essential value lies in the description of the basic structures and dynamics of personality; the delineation of the subject's mode of intellectual approach to problem-solving, his intellectual efficiency, grasp on reality and so on, on the one hand, and his affective responsiveness to environmental stimuli, his intrapsychic life, the manner in which he handles anxiety, his defenses in the face of strong emotions and of

[1]Presented at the 1949 annual meeting of the American Orthopsychiatric Association, Chicago, Illinios.

[2]From the Division of Psychology. This paper is the sixth in a series of Studies in Primary Behavior Disorders and Psychopathic Personality, under the general direction of Jacques S. Gottlieb, M.D.

paramount importance the reciprocal effects of the former upon the latter. It is this sort of approach which is too often missing in the personality investigations of the several nosological groups. We shall attempt to make such an approach in the present paper with respect to the groups under consideration.

In general, the development of primary behavior disorders in children and psychopathic personalities in adults are viewed from two standpoints: (1) that which regards these personalities as constitutionally determined and hence *genogenic*, and (2) that which regards these personalities as based in the early experiences of the individual and hence *psychogenic*. It would also appear, however, that a third viewpoint, one combining both the genogenic and the psychogenic, might also be considered. Studies by means of che electroencephalogram[2, 3, 7] have served to demonstrate the inheritance of pathophysiological processes in similar groups. In the present investigation it is proposed that the more psychological processes of personality be investigated in order to shed additional light on the genetic problem.

Method

The study employs a group of 46 children and adolescents diagnosed as primary behavior disorder and their parents. The definition of primary behavior disorder employed has been described elsewhere[3]. The group of patients includes both sexes and range in age from 8 through 22 years with the mean age being 14 years. Thirty-four were males, 12 were females. The actual distribution of ages may be seen in Table 1. Intelligence as a factor was controlled to the extent that no patients found to be feebleminded were included in the study. The range in I. Q. is from 70 to 123.

For purposes of this study each patient referred to the Out Patient Clinic was accompanied by both parents and all three members of the family were administered the Rorschach method by one of the writers. As in previous studies in this series, criteria for inclusion in the present investigation were a composite record of the psychiatric, physical, neurologic, laboratory, and psychometric findings and a detailed social history. No cases were included in which examinations indicated a psychotic or neurotic condition, convulsive disorders including epilepsy, behavior disorders with somatic disease or mental deficiency. Each record was scored with reference to Hertz' Frequency Tables[4], summarized on the psychogram, and a brief interpretation written for the record.

The literature to date reveals two general approaches which have been made in using Rorschach data experimentally: one of these is the statistical manipulation of the "raw" Rorschach scores and percentages, and the other is the attempt to fit interpretative write-ups into a rating scale system for analysis. Neither of these is entirely satisfactory since the former often leads to equivocal results, e.g., obtaining a significant difference between the number of whole responses while the percentage of whole responses does not show such a difference; and the latter is too subject to the differences in language used in writing reports by different writers, e.g., one may always mention the Rorschach evidence for the nature of "intellectual control" while

Table I
Distribution of the Patient Group by Age Levels

Age	N	Age	N	Age	N
8	3	13	4	18	0
9	2	14	5	19	1
10	2	15	7	20	1
11	2	16	7	21	0
12	7	17	4	22	1

Total N = 46

another may only mention it when it is pathologically rigid, etc. In order to overcome these defects the writers devised a rating scale which, it was felt, had the advantages of being quantified and hence manipulable statistically and in addition was on the interpretative level. In addition to these considerations the definitions of the several variables were couched in Rorschach terminology. A sample of several of the variables is given in Fig. 1.

FIG. 1 THE RORSCHACH RATING SCALE INSTRUCTIONS AND TWO OF THE VARIABLES RATED.

The variables listed below are to be rated on a scale from 0 to 5. Each rating may be referred to a section of the normal curve. The ratings should be made as though the variables were on a continuum from the aspect mentioned first to the one mentioned second. Thus, Productivity of Associations is rated on a continuum from "paucity" to "very productive" as follows: "0" extreme paucity, "1" moderate paucity, "2" below average productivity, "3" above average productivity, "4" quite productive, and "5" very productive.

The Variables	Rating Scale					
	0	1	2	3	4	5
1) Paucity-Productivity	|——	——	——	——	——	——|
14) Emotional depth: Lacking-Labile	|——	——	——	——	——	——|

It may be seen that the rating scale is so devised as to exclude the possibility of "dumping" doubtful cases in the middle since there are an even number of intervals. The instructions accompanying the rating form suggest referring the rating to a section of the normal frequency curve with the reference population being that of the normal population. Twenty variables were selected for study. These variables fall into three broad and somewhat overlapping categories: (1) 12 essentially intellectual attributes, (2) 7 essentially emotional attributes, and (3) a single rating on a continuum from extreme pathology to excellent adjustment called "Overall Pathology."

The intellectual variables included were the following:

1) Paucity vs. productivity of associations defined as the facility with which the subjects were able to call up associations to the ink-blot designs.

2) Emphasis on generalization in approaching problem-solving defined as the tendency and ability to organize and synthesize details into meaningful wholes.

3) Emphasis on thinking which is predominately concrete defined as the tendency and ability to attend to the more obvious details of the problem.

4) Emphasis on the finer nuances in problem-solving defined operationally as the tendency and ability to attend to the more precise and less apparent aspects of the problem.

5) Self-will determination defined as the tendency to "go beyond" the strict instructions in the Rorschach experiment and to employ the white spaces as stimuli.

6) Rigidity vs. refinement of intellectual control.

7) Accuracy of thinking based upon accuracy of perceptions insofar as their appositeness to reality is concerned.

8) Efficiency of intellectual functioning based upon the ratio of organizational abilities to creative capacities.

9) Stereotypy in the thinking processes.

10) Conventionality of thinking defined in terms of Rorschach norms and frequency values.

11) Originality of thinking defined normatively.

12) Originality of thinking defined qualitatively.

The more affective variables included the following:

1) Personality type defined as the balance between extratensive and introversive tendencies.

2) Inner maturity vs. immaturity defined as the level of maturity of the phantasy life.

3) Emotional balance vs. imbalance defined as the balance between controlled and more uncontrolled emotional reactivity.

4) Depth of the emotional life based upon amount of affective participation in responding to affectively toned stimuli.

5) Disturbance in emotional responsiveness defined as the extent to which behavior reflected a neurotic emotional disturbance reaction.

6) Quality of the mood life on a continuum from extreme dysphoria to extreme euphoria.

7) Sensitivity and tact in reacting to environmental stimuli.

In addition, a single rating was made based upon the clinician's overall impression of the Rorschach protocol and the psychogram as to its general indications of the presence of extreme pathology on the one hand or excellent adjustment on the other.

In rating each Rorschach record, reference was constantly made to available norms, care being taken that no shifting in standards took place during the total rating process. The forms were rated randomly excluding the possibility of bias entering into the rating if the members of families were rated together. Re-rating checks were made to assure consistency in standards and accuracy of ratings insofar as the "direction" of the scale was concerned.

After trying various statistical methods experimentally, analysis of variance proved to be the method of choice, since we are interested here in the general question of whether or not the differences obtained among the means of the members of the families on a given variable are significant of true differences, or may be explained in terms of chance fluctuations in random sampling. In other words, we want to test the hypothesis that among the members of the family the mean ratings are random samples from the same population. The application of this technique follows that described by Lindquist[9]. Obviously it follows that if this hypothesis is rejected it is required to take a second step, i.e. to apply a test of the significance of differences between the individual means to determine which accounts for the excessive variance.

Results

The results of the analysis of intrafamilial Rorschach ratings are to be seen in Tables 2 and 3. Table 2 presents the data obtained by means of analysis of variance in the case of each of the more

Table 2

Results of Analysis of Variance Applied to the Members of Families
(Intellectual Variables)

Variable	Mean Ratings			"F"
	Pt.	Fa.	Mo.	
1. Productivity	1.39	0.98	1.26	2.46
2. Generalization	3.07	3.61	3.59	2.49
3. Concrete	2.67	2.13	2.26	*3.61**
4. Precision	1.16	1.35	1.26	**
5. Self Will	0.61	0.50	0.70	
6. Rigid control	3.07	2.72	2.57	1.49
7. Accuracy	2.11	2.50	2.91	*3.69*
8. Efficiency	1.22	1.02	0.93	
9. Sterotypy	3.61	3.22	3.54	
10. Conventionality	1.54	0.98	1.88	*3.52*
11. Quan. Original	0.63	0.59	0.35	
12. Qual. Original	2.18	1.68	1.89	

* Italicized "F" values are significant at less than the 5% level.
** "F" values are not computed when less than unity.

TABLE 3
RESULTS OF ANALYSIS OF VARIANCE APPLIED TO MEMBERS OF FAMILIES.
(Emotional variables and "overall pathology")

Variable	Mean Ratings			"F"
	Pt.	Fa.	Mo.	
1. Erlebnistypus	2.54	2.76	3.12	
2. Inner maturity	0.71	1.09	0.95	1.80
3. Emotional balance	0.80	0.85	0.80	
4. Emotional depth	1.85	2.00	2.46	1.48
5. Disturbance	3.04	3.00	2.52	1.39
6. Mood quality	1.46	1.63	1.72	1.56
7. Sensitivity	1.09	1.50	2.17	*8.95*
8. Pathology	4.67	4.20	4.26	*10.94*

TABLE 4
SIGNIFICANCE OF DIFFERENCES BETWEEN MEANS AMONG FAMILY MEMBERS

Variables	Mean Differences		"t"
3. Concrete	Pt-Fa	0.54	*2.27**
	Pt-Mo	0.37	1.58
	Fa-Mo	0.17	0.71
7. Accuracy	Pt-Fa	0.39	1.28
	Pt-Mo	0.80	*2.62*
	Fa-Mo	0.41	1.34
10. Conventionality	Pt-Fa	0.30	1.16
	Pt-Mo	0.35	1.36
	Fa-Mo	0.65	*2.52*
17. Sensitivity	Pt-Fa	0.41	1.60
	Pt-Mo	1.08	*4.22*
	Fa-Mo	0.67	*2.62*
20. Overall pathology	Pt-Fa	0.47	*4.23*
	Pt-Mo	0.41	*3.69*
	Fa-Mo	0.06	0.54

* Italicized "t" values are significant at less than the 5% level of confidence.

intellectual variables; Table 3 presents the data on the more emotional factors and the overall ratings on pathology. In all but five of the variables the F value is such that we retain our hypothesis that insofar as each variable is concerned the offspring and their parents come from the same population, i.e. that any difference in variance is primarily due to chance fluctuations and do not represent true differences.

In the remaining five instances it would appear that the obtained differences are true ones and that on those variables one or more of the members deviates to a significant degree. In these cases a further step, that of computing the significance of differences between the three sets of means, was taken. The results of this analysis are presented in Table 4.

The general results are strikingly consistent, i.e., in 15 of the 20 variables the patients are similar to both of their parents. In two cases they differ significantly from their mothers and are more like their fathers while they are more like their mothers in only one instance. In the remaining variable the parents are the ones who are alike and differ from their offspring to a significant degree.

These results may also be viewed from the standpoint of the relation of the several central tendencies to what is normally expected in a healthy population. In general, these data may be seen graphically in Figs. 2 and 3, with the variables indicated along the baseline and the normally expected range indicated as falling between ratings of "2" and "3." It should be noted that not all of the variables may be depicted

thus, since in some instances a rating of "0" might indicate normalcy such as in the case of "Emotional balance v. Emotional imbalance" or "Mature intrapsychic life v. Immature intrapsychic life," nevertheless, the direction of the trends is dramatic. Again, some of the variables do not represent the same kind of distribution; for example Variable 10, Erlebnistypus, is an attempt to catalogue by ratings a personality characteristic not regarded as on a continuum. By way of illustration, this variable was rated "0" in the case of a Rorschach protocol indicating extreme introversiveness and "5" in the case of extreme extratensiveness with ambiequality and coarctation being rated "2" or "3" depending upon nuances in the psychogram regarded as throwing either of these designations one way or the other. This variable does not appear in the graph, therefore. The mean ratings of those variables represented in the profile are arranged so that a high rating signifies an excess and a low rating a deficit of the particular personality characteristic. Thus, a mean rating of 3.61 on variable 9 indicates excessive stereotypy of thinking; a rating of 1.09 on variable 17 indicates a lack of social tact and sensitivity, etc.

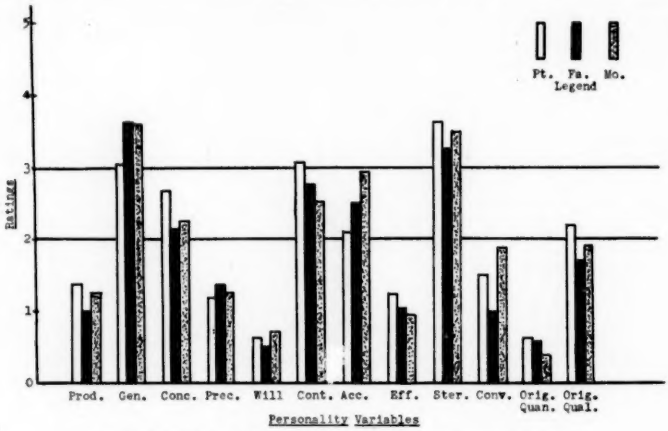


Fig. 2 Configuration of Intellectual Variables Among Family Members

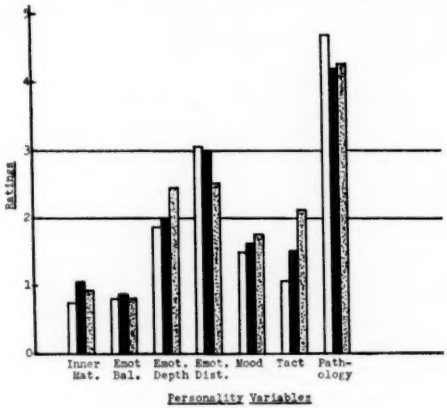


Fig. 3 Configuration of Emotional and Pathology Variables Among the Family Members

Discussion

In discussing the nature of personality as seen by means of any cross-sectional technique we are confronted with the general questions: "How did this personality get like this? What were the factors which, working together, operated to produce such a personality structure?" The evidence amassed by clinical psychiatry, psychoanalysis, cultural anthropology, and social work lead us to make the presumption that personality is produced by the familial and societal environments in which the child grows up. Of these it is also good doctrine that the most important of these influences in shaping and conditioning personality characteristics are the interpersonal relationships of the child and his parents. This is true no matter what the inherited endowment of the individual may be. Given a defective inheritance, much can be done to produce a relatively stable adjustment in a child by proper early socialization conditioning; likewise given a perfect organism, maladjustment can be produced by inferior conditioning. This leads directly to our first hypothesis which may be stated as follows: If early socialization in the form of parental conditioning in addition to genogenic factors are determinants in personality development, there should exist a communality of personality attributes among children and their parents.

The data as presented in Tables 2 and 3 and in Figures 2 and 3 point in precisely this direction. The communality is, indeed, almost dramatic. Our data reveal that on 15 of twenty personality characteristics, what slight differences do exist are due to random sampling errors and are not indicative of true differences. Stated differently, on fifteen attributes the children and their parents come from the same population on each variable. This similarity of personality structure is even more marked when it is considered that in those cases in which the patients are not like both parents, *they are always similar to one of them*, with the exception of "Overall Pathology." Thus, in the instance of concreteness of thinking the patient group differs significantly from their fathers, but are similar to their mothers. In accuracy of thinking this is reversed in that the children appear to be similar to their fathers and to differ significantly from their mothers.

"Overall Pathology" appears at first glance to be a special case since here the patients differ to a significant degree from both parents. In this characteristic it is the parents who are most alike. It is felt that this is a most significant finding and, hence, deserves special discussion.

Since the evidence favors the retention of our hypothesis that there is a communality of personality attributes among the members of the family group which leads in turn to corroborating the evidence in favor of the effects of early parental conditioning and /or inherited tendencies in producing personality, the explanation for the special case of "Overall Pathology" seems clear. It must be due to a cumulative effect of maladjustive personality characteristics obtained from parents who possess similar traits to varying degrees. That is to say, if both parents are prone to make vague, easy generalizations based upon insufficient evidence, it seems reasonable to assume that their offspring will not only do likewise but may well show such an intellectual characteristic to a more marked degree than either parent taken separately. We might refer to such an effect as a *cumulative reinforcement effect* in personality development.

Since the children-patients are extremely pathological on this rating variable (and so are their parents!), we should expect to find in the general structure of the intrafamilial personality configuration a patterning of attributes which deviates from expected norms for the normal, healthy individual. This is obviously the case as presented graphically in Figures 2 and 3.

General productivity is markedly decreased from normal expectations and reflects in the family group a lack of motivation and an inability to call up a wealth of associations either due to a lack of energy or to a paucity of associational engrams. In the present case it would appear to be more a case of the latter in combination with a relatively poor ability to establish good interper-

sonal relationships and, hence, of poor motivation. The intellectual life is marked by excessive generalization in the thinking processes. While this is frequently a pattern found among groups with superior intellectual capacities, artists, philosophers and the like, a qualitative investigation of the generalizations of the present groups reveals them to be vague, easy ones. Thus, of the Rorschach patterns which indicate this characteristic, an average of 72 percent were vague, easy or inaccurately perceived. Consequently, this would seem to indicate a tendency to generalize on the basis of superficialities and upon inadequate stimuli as opposed to a more carefully organized approach representative of a high level of abstract thinking. As groups the family members are capable of giving adequate consideration to the practical considerations of everyday living and to combine with their loose generalizations an average amount of concreteness. Attention to the finer nuances in problem-solving is below normal expectations undoubtedly due to the aforementioned approach by generalization. Self-will determination is markedly lacking in all the members of the family groups. This latter factor in the Rorschach method is based in part upon the responses involving white spaces. Our interpretation of the extreme lack of such percepts in the present groups is based upon Beck's statement[a] to the effect that "in the healthy and intelligent individual, it stands for a resolution and perseverance that carry him through, in spite of obstacles and disappointments, to a well understood objective." Our families quite obviously are lacking this sort of resolution and perseverance.

The literature on psychopathy reveals a general agreement that there is no correlation between the presence of the syndrome and intelligence; psychopaths may be dull on the one hand or decidedly superior individuals on the other hand. Our findings corroborate this concept since the family members tend to fall within the average range in accuracy of thinking. By inference this has the further connotation of an adequate grasp on reality and serves to exclude the possibility of psychosis at the time of examination. While these appear to be individuals of average general mental abilities, there is a marked inefficiency of intellectual functioning due primarily to a strong drive to impress others with their abilities which, since they are lacking in creative capacities, lends a hollow quality to their intellectual life. This is further emphasized by the marked stereotypy evident in their thinking processes. There is no evidence in general of neurotic rigidity and constriction in the conscious, voluntary intellectual control pattern; here the control factors fall almost entirely within the range of normal. Commonality of thinking, an index to intellectual and social adaptability, is clearly below normal and serves to focus the spotlight upon the basic defect underlying the thinking of this type of personality. That is to say, they reveal an inaptitude for participating in the conventional or common thinking of society. This may certainly be interpreted at least as the uncooperative attitude and is not incompatible with delinquency. As might be expected from all of the foregoing, these subjects reveal little or no originality in their thinking which places additional meaning on the findings indicative of inflexibility and lack of creativity. When it is expressed, however, it is not particularly deviant insofar as quality is concerned.

In the affective sphere, inner maturity, that is a mature state of development of the inner life of fantasy and creativity, is decreased from normal and there is an excess of the more immature which is suggestive of a fixation at a more infantile level of psychic development. Their responsiveness to emotionally toned stimuli emanating from their surroundings is markedly imbalanced, labile and often unpredictable, reflecting a fairly high degree of egocentricity in the personality configuration. With the exception of the mothers, emotionality is inclined to be superficial and lacking in depth; the fathers tending to fall at the lower limit of normal depth of feeling while the children are distinctly below average. Reactions which may be characterized as neurotic emotional disturbance are minimal and at most are

only suggestive since the children and their fathers show disturbance reactions which are at the top limit of normal; the mothers seem entirely to fall within normal limits. The general mood of all three family members is on the dysphoric side and suggests an attitude akin to looking "through a glass darkly" in their view of the world. Tact and sensitivity is decreased among the children and their fathers, the mothers revealing significantly greater tact which also falls within normal limits of expectation. The personality type tends to be coarcted for all of the family members although the mothers reveal a slight tendency in the direction of extratensiveness which is in keeping with their more nearly normal depth of emotionality and relative absence of neurotic emotional disturbance.

In summary, then, the general characteristics of personality attributes bear out our hypothesis that since the fathers and mothers do reveal a series of psychopathological personality traits to varying degrees, the children usually showing an exaggeration of these personality deficiencies on the individual trait ratings, the reason for the significantly greater pathology among the children is probably due to what we have termed a *cumulative reinforcement effect* in personality development.

The foregoing analysis still leaves at least one problem unexplained in our data if our basic premise of the importance of early conditioning is to be maintained. That is the question of why the mothers and fathers are so strikingly similar. Obviously they could not have had such a basic effect on each other during the socialization experiences of marriage (they do not strike us as that type of personality). We are left with the possibility that each marriage partner must have had relatively similar early parental and environmental conditioning experiences as children and perhaps similar endowments and that somehow a selective factor operated in their choice of each other as mates.

Various investigators of the factors involved in the selection of marriage partners have reported evidence that homogamy rather than heterogamy is the rule. The types of psychological characteristics studied can be classified as intellectual, temperamental, or attitudinal. With respect to intellectual factors, the correlations in intelligence between husband and wife are positive. Jones[6] found a correlation of .598 in similarity of intelligence as measured by the Army Alpha test, for example; others have reported correlations ranging from .42 to .56.

Correlations in the area of attitudes range from .38 to .76. Terman[13] obtained answers from 792 couples concerning congeniality of tastes and interests. On items tapping attitudes toward social conventions, recreation, religion, demonstrations of affection, friends, caring for children, and philosophy of life, agreements ranged from 93 per cent to 96 per cent when all essential agreements are combined against essential disagreements.

Penrose[10] states that for "one mental hospital in Canada, the frequency of both husband and wife being admitted with a psychotic illness was much greater than chance." From this he concludes that assortative mating does exist with respect to traits which form part of the background of mental disease. Slater[12] and Woodside[14] interviewed and administered tests of temperamental traits to a group of neurotic soldiers, a group of soldiers from the general wards of a hospital and the wives of both groups. Of the first fifty couples in each group, the incidence of neurosis was found to be twice as high among the wives of the neurotic soldiers as among the wives of the non-neurotic soldiers.

In summarizing the research in the area of mental resemblances between husbands and wives, Richardson[11] makes the following generalization: "The correlations are higher in the intellectual and attitudinal traits than in traits of temperament. None of the correlations, even in temperament, are reliably negative, as the theory of the attraction of opposites would require."

Summary and Conclusions

Forty-six children and adolescents, selected because they fit the classification of primary behavior disorder in children or psychopathic personality,

and their parents were studied by means of a specially devised rating scale in the rating of which Rorschach protocols were used as the sole source of information about their personality structures. The data obtained from the rating scales were analyzed by means of analysis of variance techniques in order to test the hypothesis that the individual family members are homogeneous on the several variables studied. A second hypothesis derived from the data concerned the question of whether or not the early socialization conditioning and inherited characteristics of children of pathological parents had a cumulative conditioning effect upon the nature of personality development in the children. In the course of the latter discussion the structure of the intrafamilial personality configuration was described in detail. Evidence is presented to account in part for the similarity between the parents based upon a selective factor in the selection of marriage partners.

On the basis of the present data, the following conclusions may be drawn:

1) On fifteen of the twenty personality variables studied, there appears to be evidence that significant homogeneity exists among all three members of the family groups. On four other variables it has been shown that homogeneity exists between the children and at least one parent.

2) Since in one variable, the rating on "Overall Pathology," the children differ significantly from both parents, all three members of the family being extremely pathological however, evidence is presented to support the contention that this greater pathology in the children is due to a *cumulative reinforcement effect* in the personality development of children by parents possessing pathological personality traits to varying degrees.

3) The general structure of the personality of subjects diagnosed as behavior disorder or psychopathic personality is described based on quantitative evidence. It is felt that this description fits closely that prevalent in much of the literature based on more informal kinds of evidence.

4) In the light of the latter result, it would appear that the study also sheds light on the validity with which personality attributes may be discerned by means of the Rorschach method.

These conclusions are presented within the framework of a psychological theory which regards the genesis of personality attributes as based in multiple sources. Personality development involves a constellation of factors of which the psychogenic and genogenic are only two. Projected research plans call for studies aimed at the more physiogenic factors, factors based in social elements, and combined research aimed at eventually partialling out psychogenic and genogenic factors. It is felt that the presence of pathology in certain of the characteristics revealed by the Rorschach method are psychogenic in origin, some are quite possibly genogenic. The Rorschach technique, like most psychological tools, cannot reflect the total etiology; research carried on jointly with the Rorschach and, for example, the electroencephalograph may assist in the clarification of some of these problems.

Bibliography

1. Beck, S. J. *Rorschach's Test, Vol. II: A Variety of Personality Pictures*, New York: Grune and Stratton, 1946.
2. Gottlieb, J. S., Ashby, M. C. and Knott, J. R. Primary behavior disorders and psychopathic personality: I. Correlation of the electroencephalogram with family history and antecedent illness or injury. *Arch. Neurol. Psychiat.*, 1946, 56, 381-400.
3. Gottlieb, J. S., Knott, J. R. and Ashby, M. C. Electroencephalographic evaluation of primary behavior disorders in children. *Arch. Neurol. Psychiat.*, 1945, 53, 138-143.
4. Hertz, M. R. *Frequency Tables to be used in Scoring the Rorschach Ink blot Test.* Cleveland: Western Reserve University, 1946.
5. Heuser, K. D. The psychopathic personality: The Rorschach patterns of 28 cases. *Amer. J. Psychiat.*, 1946, 103, 105-112.
6. Jones, H. E. A first study of parent-child resemblance. *27th Yearbook of the National Society for the Study of Education.* 1928, Part I, 61-72.
7. Knott, J. R. and Gottlieb, J. S. The electroencephalogram in psychopathic personality. *Psychosomatic Medicine*, 1943, 5, 139-142.
8. Linder, R. M. The Rorschach test and the diagnosis of psychopathic personality. *J. Crim. Psychopathol.*, 1943, 1, 69 ff.

9. LINDQUIST, E. F. *Statistical Analysis in Educational Research*, New York: Houghton Mifflin, 1940.
10. PENROSE, L. S. Psychiatric Quarterly Supplement, 1944.
11. RICHARDSON, H. M. Studies of mental resemblances between husbands and wives and between friends. *Psychol. Bull.*, 1939, 36, 104-120.
12. SLATER, ELIOT. An investigation into assortative mating. *Eugenics Review*, 1946, 38, 27-28.
13. TERMAN, L. M. *Psychological factors in marital happiness*. New York: McGraw-Hill, 1938.
14. WOODSIDE, M. Courtship and mating in an urban community. *Eugenics Review*, 1946, 38, 29-39.

OBJECTIVE STUDIES OF CHILDREN'S DRAWINGS OF HUMAN FIGURES

I. SEX AWARENESS AND SOCIO-ECONOMIC LEVEL

ARTHUR WEIDER, PH.D. AND PAUL A. NOLLER, B.S.

University of Louisville School of Medicine

INTRODUCTION

The ability to differentiate characteristics between the sexes cosmetically, structurally, anatomically and psychologically is of decided importance in the sex adjustment of the individual. Awareness of these sex characteristics is a step in building healthy and mature attitudes and in eliminating undesirable ones. It is, therefore, advisable to determine what conditions favor or contribute to the development of this capacity in an individual. Since the socio-economic level of a person's environment is a factor in general personality development, it has been selected as the determinant to be studied in respect to a facet of personality in this paper.

In studying a level of genetic development in children, it is important to obtain facts supplied by the children themselves, if possible. To do this, one must use methods for meeting children on their own level of expression. Projective methods, generally, and drawings of children, specifically, have been successfully used to study personality development and deviations (1, 5, 6, 7, 8). Even at an early age, children become aware of the cultural taboos in noting sex differences so that this study might be considered a measure of the cultural restrictions in expressing awareness of sex differentiation between socio-economic levels.

Our present interest originated during use of the Goodenough draw-a-man test with children for the measurement of intellectual level[1]. Persistent evidences of the possibility that projective material was frequently offered for further study beyond that of evaluation of intelligence, finally convinced us of the need to concentrate on one area of the problem. Goodenough made observations on the sexual differences between the man drawing produced by boys and girls, but confined this to the differences shown as related to the femininity or masculinity of the individual child projected on the man drawing[2]. Morris[6] has offered methodological and some normative data in appraising drawings of children. From a perusal of the children's drawings and after studying the background and identifying material available, it became apparent to us that other variables were operating. To this end we decided to explore the socio-economic level of the child. On the basis of these observations, the hypothesis was made concerning the differences in the ways that children in different socio-economic levels express awareness of sex characteristics pictorially and verbally.

We have investigated the problem from a different approach and developed a procedure more along the lines of a draw-a-person test than is confined by the draw-a-man test of Goodenough. Since verbal responses are more symbolic and more subject to conscious manipulations than graphic projections[5], *p. 21*, it was found that a substantial amount of material expressed graphically was

ignored verbally. Because of this, individual interviews were used to substantiate the material pictorially presented by the children in the attempt to get as much objective information from them as possible. Moreover, our technique was used in a group procedure in order to facilitate the acquisition of a larger sample.

The authors assume there is a difference in expressing awareness of sex characteristics through the use of drawings and by interviews between children of different socio-economic levels. We investigated the objective expression of the awareness of sex differences according to the following criteria:

(1) The number of sex characteristics actually drawn, e.g., hair, basic clothing, physical differences, cosmetic - facial differences, specific items of clothing, accessories, etc.

(2) "Identification", that is, the sex of the first drawing in relation to the sex of the subject and in relation to the sex of the larger drawing.

(3) The difference between covert and overt sex characteristics, that is, difference between what the subject actually drew as against what he revealed in the interview as being significant in determining the sex of the drawings.

A later paper will be concerned with normative date for this group, draw-a-person technique to be used with children[8]. These norms will be established separately for boys' and girls' drawings in respect to the characteristics that are influenced by sex differences of the subject. Where we find no difference, norms for children in general will be determined. In other words, this present study may therefore be considered a pilot study for a more comprehensive investigation of the objective and projective properties of the draw-a-person technique for children.

Methods and Procedures

The classifications used in a Louisville study[4] for determining socio-economic level on which class differences can be based give a sufficient number of social levels for investigation. These classifications are based on the following criteria: average year schooling, average rental values, percent dwellings with more than 1.5 persons per room, etc. Differences in socio-economic levels were based on the residential location of the children. To establish an objective basis for ascertaining socio-economic differences, the neighborhood areas as set up by the Planning and Zoning Commission of the city of Louisville in association with the Department of Sociology of the University of Louisville as determined by various socio-economic characteristics of the 1940 census was used. These neighborhoods were divided into five classifications: (A) upper, (B) upper middle, (C) middle, (D) lower middle, (E) lower. Only those children living in areas classified as (A) upper, (C) middle, and (E) lower were used to obtain an adequate sample of three sharply defined levels of socio-economic classes in order to avoid over-lapping.

Schools were selected in areas (A), (C), or (E) so that the school would draw enrollment from at least one of the levels. At least two schools, and sometimes three or four were used to supply children from each socio-economic level so that no level would be represented by only one individual school or the prevailing philosophy (bias) of that school. In each school used, all third grades participated so that no one individual personality of any teacher would effect the results obtained from that particular school.

Prior to administration, class lists (supplied by individual teachers according to mimeographed instructions) from the schools were obtained with the following information: child's full name, birth date, intelligence quotient made on the Otis Quick Scoring Mental Ability Test. Using the above information home locations were plotted and children living in areas (B) and (D) were deleted from the list. IQ distributions were made for all three levels and only those children of average intelligence (with IQs between 90 and 110 inclusive) were retained on the list. Children for inter-

viewing were thus designated prior to administration of the tests. Each child was given two sheets of paper (8½" by 11") and asked to put one sheet in his desk and to have the other sheet and a sharpened pencil ready for use. These instructions were then given: "Draw a person, draw a whole person; take your time and draw the best person you can." They were further assured that this was not to see how well they could draw but to see how they could draw a person. This task took approximately 10 minutes, but the class waited until all of the children had completed their drawing and labeled it (1—full name of child). Drawing 1 was then put away and the second sheet brought forth. The children were then told: "There are two kinds of persons—some of you drew a boy or a man, and some of you drew a girl or a woman. This time you are to draw the other kind of person. Those of you who drew a boy or a man, will now draw a girl or a woman, and those of you who drew a girl or a woman will now a boy or a man." They then identified this drawing as (2—full name of child) approximately 10 minutes after initiation of second drawing. Both drawings were collected after the children checked their identifying data. The children were then told that the examiner would now talk to some of the children individually about their drawings and that, as there would not be sufficient time to talk to all of them, just randomly selected children, that is, every third or fourth child would be chosen. However, those children who fulfilled the criteria for this investigation, that is: 1) were in the age group between 8 and 10; 2) had an Otis IQ of between 90 and 110; and 3) had residence in (A) or (C) or (E) socio-economic levels were actually chosen.

Interviewing was conducted privately, outside the class room, immediately after the drawing procedure and consisted of the following questions asked concerning the first drawing and then the second drawing: (1) What is this? (2) What about this drawing shows it is a boy or a girl? (3) How old is this boy (girl)? (4) Who is this boy (girl)? (5) Is it anyone you know? (6) Does he (she) have a name? (7) Specific questioning as to features which were not verbalized in answer to question 2 and reactions were noted.

Subjects

Table I shows the classification of subjects according to socio-economic levels, sex, IQ, and chronological age. It will be noted that the various groups are for all practical purposes equated according to the fundamental variables.

Table 1

Composition of Population By Sex, Age, and IQ According to Socio-Economic Levels

Socio-Economic Level		Sex of Subjects	N	CA	Mean Otis IQ	Standard Deviation
UPPER	(A	Boys	10		103.8	
	(A	Girls	24		100.6	
	(A	Total	34	8–11	101.6	5.5
MIDDLE	(C	Boys	34		99.0	
	(C	Girls	23		100.1	
	(C	Total	57	8–10	100.0	5.5
LOWER	(E	Boys	29		98.8	
	(E	Girls	33		96.6	
	(E	Total	62	8–11	98.0	5.4
ALL	(ACE	Boys	73	8–11	100.5	5.0
	(ACE	Girls	80	8–11	99.0	5.0
	(ACE	Total	153	8–11	99.0	5.0

Results

Table II shows the relationship between the sex of the subject and the sex of the first figure drawn for the various socio-economic levels, and $X^2 = 17.84$ reflects a real sex difference as the girls drew the same sex first more often than the boys. However, $X^2 = 3.79$ indicates that the differences between the economic levels could be due to chance, while $X^2 = .04$ and $X^2 = 3.82$ indicate that differences between these socio-economic levels in all probability are due to chance.

Table III shows the relationship between the sex of the subject and the larger figure drawn for various socio-economic levels. The girls drew their own sex larger more often than the boys and this difference between the sexes is reliable; however, the differences between the economic levels are probably due to chance.

Table IV shows the relationship between the socio-economic level and the overt responses verbalized and between covert characteristics drawn for the male

Table 2

The Relationship Between the Sex of the Subject and the First Figure Drawn for Various Socio-Economic Levels

Subjects	Total First Figure Drawn	Per Cent		Probability That Results Are Due To Chance
		Same Sex	Opp. Sex	
All Subjects	153	86.3	13.7	Less than
Boys	73	74.0	26.0	1 in 1,000
Girls	80	97.5	2.5	$X^2 = 17.84$
Total Boys	73	74.0	26.0	16 in 100
Econ. Level A	10	70.0	30.0	$X^2 = 3.79$
" " C	34	64.7	35.3	
" " E	29	86.2	13.8	
Total Girls	80	97.5	2.5	More than
Econ. Level A	24	95.8	4.2	9 out of 10
" " C	23	95.7	4.3	$X^2 = .04$
" " E	33	100.0	0	
Total Subjects	153	86.3	13.7	44 in 100
Econ. Level A	34	88.2	11.8	$X^2 = 3.83$
" " C	57	77.2	22.8	
" " E	62	93.5	6.5	

Table 3

The Relationship Between the Sex of the Subject and the Larger Figure Drawn for Various Socio-Economic Levels

Subjects	Total Larger Figure Drawn	Per Cent		Probability That Results Are Due To Chance
		Same Sex Drawn Larger	Opp. Sex Drawn Larger	
All Subjects	153	66.7	33.3	Less than
Boys	73	52.1	47.9	1 in 1,000
Girls	80	80.0	20.0	$X^2 = 13.41$
All Economic Levels	153	66.7	33.3	49 in 100
Level A	34	73.5	26.5	$X^2 = 1.46$
" C	57	61.4	38.6	
" E	62	67.7	32.3	

subjects. Here $X^2 = 11.16$ indicates that among boys, socio-economic level is related to the number of covert characteristics drawn for the opposite sex drawing. $X^2 = .73$ reflects that among boys, the number of verbal responses was less than the drawn characteristics regardless of socio-economic level for the opposite sex drawing. $X^2 = 8.13$ might indicate that among boys, the socio-economic level is related to the number of covert characteristics drawn for the same sex drawing. However, $X^2 = .013$ would indicate that among boys, the number of verbal responses was less than the drawn characteristics regardless of socio-economic level both for the same sex and the opposite sex.

Table V shows the relationship between the socio-economic level and the overt responses verbalized and between covert characteristics drawn for the female subjects. $X^2 = 6.05$ and $X^2 = 3.88$ show that the relationship of the socio-

TABLE 4.

THE RELATIONSHIP BETWEEN SOCIO-ECONOMIC LEVEL AND OVERT RESPONSES VERBALIZED ABD BETWEEN AND COVERT CHARACTERISTICS DRAWN

MALE SUBJECTS

Socio-Economic Level	N	Same Sex Figure: Average Per Figure	Same Sex Figure: Probability That Results Are Due to Chance	Opposite Sex Figure: Average Per Figure	Opposite Sex Figure: Probability That Results Are Due To Chance
Covert Characteristics Drawn					
All Males	73	3.89	2 in 100	3.33	6 in 1,000
Econ. Level A	10	3.10	$X^2 = 8.13$	2.70	$X^2 = 11.16$
" " C	34	3.44	(n = 2)	2.76	(n = 2)
" " E	29	4.69		4.21	
Per Cent Increase from A to E—51%				56%	
Overt Responses Verbalized					
All Males	73	2.70	99 in 100	2.52	69 in 100
Econ. Level A	10	2.70	$X^2 = .013$	2.90	$X^2 = .73$
" " C	34	2.68	(n = 2)	2.41	(n = 2)
" " E	29	2.72		2.52	

TABLE 5.

THE RELATIONSHIP BETWEEN SOCIO-ECONOMIC LEVEL AND OVERT RESPONSES VERBALIZED AND BETWEEN COVERT CHARACTERISTICS DRAWN

FEMALE SUBJECTS

Socio-Economic Level	N	Same Sex Figure: Average Per Figure	Same Sex Figure: Probability That Results Are Due To Chance	Opposite Sex Figure: Average Per Figure	Opposite Sex Figure: Probability That Results Are Due To Chance
Covert Characteristics Drawn					
All Females	80	3.86	49 in 1,000	3.35	15 in 100
Econ. Level A	24	3.04	$X^2 = 6.05$	2.75	$X^2 = 3.88$
" " C	23	4.13	(n = 2)	3.48	(n = 2)
" " E	33	4.27		3.70	
Per Cent Increase A to E 40.46				34.5	
Overt Responses Verbalized					
All Females	80	2.75	71 in 100	2.63	68 in 100
Econ. Level A	24	2.58	$X^2 = .70$	2.54	$X^2 = .79$
" " C	23	3.00	(n = 2)	2.74	(n = 2)
" " E	33	2.70		2.61	

economic level for girls is not so clearly indicated. The same sex figure is significant to the 5% level of significance but not the opposite sex drawing. $X^2 = .70$ and $X^2 = .79$ indicate that socio-economic level is not related to the number of overt responses verbalized either for the same sex drawing or the opposite sex drawing.

Discussion

Concerning the relationship between the sex of the subject and the first figure drawn, three-quarters of the boys (54 of 73) and nearly all of the girls (78 of 80) drew their own sex first. Socio-economic level apparently is unrelated to which sex would be drawn first. From this data we might infer that girls between the ages of 8 and 10 (the so-called "latency period") "identify" more with their own sex than boys between 8 and 10.

Regarding the relationship between the sex of the subject and the larger figure drawn: boys are almost evenly divided (38 of 73) as to whether they drew their own sex larger or the opposite sex larger, while girls were more inclined to draw their own sex larger (64 of 80). There is a real difference between the sexes in this respect. However, socio-economic level again was unrelated to which sex was drawn larger, all differences being due to chance. From this data, we might infer that this tendency for the girls to "identify" more with their own sex than the boys do, is further strengthened by the greater amount of space girls use to portray their own sex, while at the same time they tend to draw the male smaller. Furthermore, we found that about half the boys (36 of 73) drew full-faced figures and half (37 of 73) drew profile figures regardless of the sex of the figure drawn. Among the girls, however, three-fourths (61 of 80) drew full-faced figures and only one-fourth (19 of 80) drew profile figures for both sexes. So that, in addition, girls draw the full face, which may offer an opportunity to give more attention to facial detail. Additional data is being investigated to determine objectively at what age this "identification" and preference begins to show itself and if it is related to any of the Freudian psychosexual stages. We feel that there is some relationship to the amount of space the child uses in drawing his own sex or the opposite sex, and to whether the drawing is full-face or profile, and also to the location of the drawing on the sheet. All of these variables are being further investigated in a larger normative study[8].

Regarding the relationship between the socio-economic level and the overt responses verbalized and between covert characteristics drawn for the male subjects: socio-economic level is related to the covert characteristics drawn by the boys for both figures. For boys and girls, however, a larger number of sex characteristics were shown in the drawings of their own sex. The number of responses verbalized did not differ significantly for the 3 socio-economic levels. Thus, we might infer that concerning different socio-economic levels, boys from socio-economic level (A) compared to boys from level (E) include more sex characteristics on their drawings of their own sex compared to the number they draw for the opposite sex.

Like the boys, socio-economic level is probably related to the number of covert characteristics drawn by girls. This is clearly indicated for their own (female) sex figure, though doubtful for their male figure, and like the boys, the girls show no relationship between socio-economic level and overt responses verbalized.

Summary and Conclusions

1. Of 73 boys, ages 8 to 10, 74% drew their own sex first.
   Of 80 girls, 97% drew their own sex first.
   Statistically significant at the 1% Level of Confidence.
   However, slight differences in socio-economic level were due to chance.
2. 52% of boys drew their own sex larger.
   80% of girls drew their own sex larger.
   Statistically significant at the 1% Level of Confidence.
   Again, slight differences in socio-economic level were due to chance.
3. For boys, socio-economic level was related to the number of characteristics drawn for both figures of their own and opposite sex with more items being included for their own sex figure.

4. As we change from socio-economic level (A) upper to level (E) lower there was a statistically reliable increase in characteristics drawn.
5. However, the number of responses verbalized in interview did not differ significantly for these socio-economic levels.
6. All of the above (Nos. 3, 4, 5) were true for the girls except that there was no statistically reliable difference in the number of characteristics drawn for their opposite sex figure.

Bibliography

1. Goodenough, Florence L. *Measurement of Intelligence by Drawings.* Yonkers, World Book Co., 1926.
2. Goodenough, Florence L. Children's Drawings, in Murcheson, C., *Handbook of Child Psychology,* Worcester, Mass., 1933, pp. 480-514.
3. Knopf, Irvin J., & Richards, T. W. *The Child's Differentiation of Sex As Reflected In Drawings of The Human Figure,* paper delivered before the American Orthopsychiatric Association, Altantic City, February 23, 1950.
4. Louisville Health & Welfare Council (mimeographed report. *Participation in Recreation and Group Work Agencies,* Feb. 1946.
5. Machover, Karen. *Personality Projection in the Drawing of the Human Figure,* Springfield, Ill., Charles C. Thomas, 1949.
6. Morris, Woodrow W. *Methodological and Normative Considerations in the Use of Human Figure Drawings As a Projective Method of Personality Diagnosis,* paper delivered before the Clinical Section, American Psychological Association, Denver, September 6, 1949.
7. Springer, N. N. A Study of the Drawings of Maladjusted and Adjusted Children. *J. Genet. Psychol.,* 1941, 56, 131-138.
8. Weider, Arthur, & Noller, P. A. Objective Studies of Children's Drawings of Human Figures. II. A Normative Study (*In Press*).

PERSONALITY CHARACTERISTICS IN HYPNOTICALLY-INDUCED HOSTILITY

ROBERT M. COUNTS AND IVAN N. MENSH

Washington University School of Medicine

Introduction

Rorschach[5] early described the significance of white space selection (S) in response to the stimulus cards as "oppositional tendency"; and Beck[1], more recently, emphasized that " . . . the personality significance of white space selection includes always a nucleus of contrariness." The present experimental design represented a method for examining experimentally this relationship between Rorschach responses and other behavior of individuals in specific conflict situations involving "opposition" or "contrariness" feelings. The technique of hypnosis was selected to alter the intensity of these feelings, which we now shall term "hostility." By means of artificially induced conflict in subjects and repeated administration of the Rorschach method, the design permitted an objective measure of changes, if any, paralleled by psychiatric interview material. Thus, there was a preliminary selection of subjects, those who were selected were examined by the Rorschach method, re-examined immediately following the induction of hostility and conflict, and twice more after removal of the suggestion.

Procedure

All but one of the subjects were volunteers answering a request for research subjects. An announcement was made in several first-year psychology classes. Students interested in receiving sixty-five cents an hour as research subjects were asked to report for a preliminary interview at the campus health service.[1] They were not informed in advance that this would involve hypnosis. JG, a graduate nurse, had been used prior to this project in a demonstration of hypnosis and readily agreed to take part in the experiment.

[1] Grateful acknowledgment is made to Dr. W. B. Webb, Department of Psychology, Washington University, for his assistance in interesting the students in this research project.

Students reporting were given a brief psychiatric interview by one of the investigators (RMC). Individuals who appeared obviously hostile or severely neurotic were not accepted as subjects. Each student wrote a brief story about each of five TAT cards (6GF, 7GF, 8BM, 12M, and 18BM). Those individuals who saw sinister events in card 12M were not accepted as subjects (this card usually is referred to as the hypnosis card). It was felt that students who, in response to this card, told stories which could be construed as favorable to hypnosis, would be more readily hypnotized than those who saw sinister events in the card. This preliminary screening technique was introduced on the assumption that it reduced the time necessary to determine whether a student would be a good hypnotic subject.

Of the thirteen who reported initially, eight students (including JG) thus were selected for futher study. These were informed that the research project would involve hypnosis, then were asked what they knew about hypnosis, and any erroneous ideas (they might not awaken, etc.) were discussed. Finally, an attempt was made to hypnotize each of the eight subjects. In one subject, RD, who apparently was not suitable for hypnosis, only a light trance resulted. Three young women subjects (WC, CB, JB) were readily hypnotized, followed post-hypnotic suggestions, and had paralyses and paresthesias during the trance. However, these three did not have amnesia for the hypnotic state when so directed, a prerequisite for our study. One of these three (CB) is included in the group of five because in the first session it appeared that there was complete amnesia. Subsequently it became apparent that the subject remembered some of the suggestions. Nevertheless, behavioral changes were observed and she was included in the study. The remaining four were all good hypnotic subjects, had complete amnesia for the trance when so directed, and carried out post-hypnotic suggestions readily (up to and including post-hypnotic positive hallucinations).

The range in ages of our subjects was 19 to 24. Three of the five were males, two were females. The subjects were preparing for the medical or closely allied professions. These five were instructed to return several days later to have a Rorschach examination. All were cooperative and interested in the project. They were introduced to INM and responded to the Rorschach cards for the first time. They returned approximately one week later for an appointment with RMC. The subjects were hypnotized and the following suggestions designed to produce an artificial conflict, were presented to each subject. This technique was adapted from Erickson's method[3] for the induction of an experimental neurosis.

I am now going to recall something that happened to you several weeks ago. It is now very vague in your mind and consciously you don't remember it. However, as I recall it to you, you will picture it in your mind's eye—how it happened, and how you felt at the time. You were just getting out of your psychology class. As you were walking down the hall, Dr. Mensh came rushing out of one of the rooms. He bumped into you. It was obviously his fault as he was not looking where he was going. He was walking so fast that you were almost knocked down. As you were recovering your balance and dignity, he said very sarcastically that you were very clumsy to bump into him, that if you didn't watch what you were doing he would make trouble for you. You immediately felt very angry but realized that you could say nothing. As you walked away you felt more and more angry and thought of many sarcastic things that you might have said but didn't. Now when you awaken you will not remember this event consciously, but it will be preying on your unconscious mind. During the Rorschach test that you are about to take with Dr. Mensh you will feel very uncomfortable and quite angry with him but you will not know why. Nevertheless, you will be able to complete the series of ten cards. You will dislike him very much during the procedure but you will not know why. You will not feel angry with me and you will be hypnotized after the procedure as readily as you just have been.

The above was read to the subject twice before he or she was awakened from the trance. Immediately after this session, the patient was taken to the office of INM who presented the Rorschach stimuli for the second time. After this session the subject was again interviewed by RMC. After RMC discussed the subject's attitudes toward INM, the subject was rehypnotized and the artificial conflict removed. He was allowed to remember the artificial conflict and also was given the post-hypnotic suggestion

that he would no longer feel resentment towards INM. One week later the subject was examined for the third time, in this session without hypnosis, and instructed to return in a week for the fourth and final Rorschach test which was administered by RMC.

Clinical Results

Four of the subjects (JG, IW, WS, WC) had complete amnesia for the artificially induced conflict, and the behavioral changes were definite. CB did not have complete amnesia and her behavior showed less change. The first four mentioned were quite hostile toward INM following their second session with him (conflict present). Each employed various defense techniques in this situation, including projection, denial, and turning inward the feelings of hostility (cf. protocols).

JG was interviewed by RMC, as were the other subjects, immediately following her second session with INM. She was friendly and cordial. She was asked if anything unusual had happened during the meeting with INM and answered, "No." Next she was asked how she felt towards INM and replied, "All right." She then was asked to describe any feelings that she might have had, whether or not they seemed illogical. She started to say something and stopped. The interviewer indicated to her that he was interested in what she had been about to say. She answered that she didn't think INM was "in very good rapport" with her. Asked to elaborate, she explained that INM was not very friendly towards her, and went on to say she was quite uncomfortable with him. However, when asked why she was uncomfortable, she was unable to answer. She was next asked directly, "Did you like or dislike Dr. Mensh?" She hedged and did not answer the question. She again was asked the same question and replied that INM wasn't very friendly, and that she had been uncomfortable with him. She wanted to know if the interviewer was going to inform INM of this and was reassured of RMC'S confidence. Following this, the patient stated more definitely that she didn't like INM. She attributed her dislike to the fact that "He is a psychologist who prys into people's personalities." She remarked that she had hurried through the Rorschach situation, and when INM found it difficult to keep up with her responses she neither felt particularly sorry for him nor had she slowed down.

After this discussion, JG was rehypnotized and asked how she felt towards INM. She replied that she disliked him. To the question, "How much?" she responded, "Very much." When asked why, she didn't know. Told that she could remember the artificial conflict, she readily recognized that the events described in it had not actually happened. Reawakened without amnesia, she realized that she had been angry and uncomfortable with INM and wanted to talk with him. She was fearful that he would dislike her for the way she had acted during the Rorschach session.

Following is a verbatim interview with a second subject, WC, immediately after the Rorschach session with induced conflict. Throughout this interview the subject seemed to be preoccupied, answered slowly and carefully, and spent most of the interview staring out of a window.

RMC: How did things go?
WC : O.K., I guess.
RMC: You have some doubts?
WC : That was just an expression. He just gave me the Rorschach again.
RMC: How was your session with him?
WC : As far as I know it was O.K.
RMC: You seem to be preoccupied today.
WC : I felt sort of strange with him.
RMC: Why?
WC : He sort of annoyed me. It wasn't anything he said.
RMC: How did he annoy you?
WC : I haven't figured it out. It was like I was mad at him, yet I wasn't. It was a strange feeling.
RMC: What sort of feelings did you have with him?
WC : I don't exactly know. It was an antagonistic feeling. I don't know why. I never felt that way before.
RMC: Could you tell me a bit more about your feelings?
WC : No. I can't understand it, as nice as he always was. I guess its something in me.
RMC: Something in you?
WC : Yeah. I didn't feel that way until I got in there. I practically fell over a chair when I went in. I'm sure that had nothing to do with it.

RMC: Do you mean that falling over a chair might have made you angry?
WC : No! I usually don't get mad at such little things. I haven't that much of a temper. (The subject sat and thought for a minute or two.)
RMC: What are you thinking about, W?
WC : I don't know. It was a strange feeling, a heavy feeling.
RMC: A heavy feeling?
WC : Yes, like anticipation, when you get the feeling something is going to happen.
RMC: Anxiety?
WC : Yes, I suppose so, a feeling of anxiety. I don't know what that would have to do with Dr. Mensh.
RMC: What did you think of Dr. Mensh today?
WC : The question isn't what I felt, but why I had those feelings toward him.
RMC: What sort of feelings toward him?
WC : Sort of annoyance, an annoyance I don't see any basis for. Sort of a dunce of a character, aren't I?
RMC: What do you mean?
WC : I feel very stupid, sort of irresponsible.
RMC: Why is that?
WC : Because I felt that way.
RMC: Felt what way?
WC : That strange feeling towards him today.
RMC: You seem to be ashamed of that feeling.
WC : I'm not particularly proud of it, a feeling without cause or without apparent cause.
RMC: Without apparent cause?
WC : He's never done anything against me.
RMC: Can you tell me some of the feelings you had during the hour?
WC : I had a feeling of resentment towards him and because I had that feeling it put me in a suppressed or depressed mood.
RMC: Were you actually depressed?
WC : Sort of heavy feeling.
RMC: Do you mean anxious?
WC : Possibly. Should I have had any such feelings?
RMC: What do you mean?
WC : It wouldn't surprise me.
RMC: Why?
WC : Yes, from talking with you.
RMC: From talking with me?
WC : Before I went down there (to see Dr. Mensh).
RMC: What do you mean?
WC : I know you were going to hypnotize me and maybe you gave me that suggestion.
RMC: Do you remember my giving you such a suggestion?
WC : No.

The subject was rehypnotized, allowed to remember the artificial conflict, then awakened. He then was asked to describe more fully his feelings toward INM. He attributed his feelings of hostility to resentment against being tested with the same cards again, and because he was questioned about his responses. Asked if he usually became angry about such things, he said that normally he didn't get angry over such minutiae.

Subjects IW and WS had reactions very similar to those of JG and WC. CB, however, did not have complete amnesia for the conflict. Moreover, she was able to remember that a post-hypnotic suggestion had been for her to be angry with INM. It was interesting that although she did not have appreciable amnesia during the post-conflict session with INM, she felt quite tense and anxious. This was in contrast to other sessions when she felt at ease.

Rorschach Data

These data were analyzed along two dimensions—response characteristic and experimental condition. The former consisted of five areas which were considered most relevant to the experimental situations, and the conditions were the four outlined above—the pre-induction session in which the Rorschach method was used by INM to study the personality characteristics of the subject, the hostility-induced session, the post-hypnotic session, and the post-hypnotic session in which RMC administered the Rorschach test to the subject. Table I summarizes these data. Because of the small number of cases, the median as well as mean values are presented, together with the range of variation. Although only five subjects were studied, the mean and median values for the several categories are relatively near one another except for RIII. In general, then, the distributions are not markedly skewed.

It can be seen that, although the mean per cent of S responses in the total increased by 50 per cent from 1.8% to 2.7% after the initial Rorschach examination, the level did not change in the three succeeding sessions. These data indicate that the changes in hostility level noted in the psychiatric interview material were not reflected in the Rorschach responses. The absolute number of responses do, however, show the

TABLE I. DATA ANALYZED ACCORDING TO 1) RORSCHACH CHARACTERISTIC AND 2) EXPERIMENTAL CONDITION RORSCHACH RESPONSE CHARACTERISTIC

			R	F+	M	C	C(Wt.)	FC Int.	S
E	I	M	77.4	[1](57.4)34.0	(0.8) 5.2	(2.6)9.6	9.5	.86	(1.8) 7.0
x		Mdn	76.0	38.0	6.0	9.0	9.0	.92	6.0
p		Rge	22–116	14–45	0–10	2–18	1–20	.50–1.15	1–14
e									
r.	II	M	97.0	(52.4)38.4	(1.6) 7.8	(1.8) 9.0	8.9	.89	(2.7)12.8
		Mdn	98.0	39.0	6.0	10.0	10.0	1.00	10.0
c		Rge	19–154	12–60	0–18	1–15	5–16	.50–1.07	0–32
o									
n	III	M	101.2	(51.7)39.0	(1.6) 8.4	(2.4)11.6	10.3	1.43	(2.6)12.6
d		Mdn	123.0	38.0	6.0	12.0	11.0	.92	12.0
i		Rge	26–154	16–53	1–18	3–17	2–16	.67–.94	1–28
t									
i	IV	M	84.8	(52.3)33.6	(1.6) 7.4	(2.1) 9.0	7.4	.79	(2.6)10.6
o		Mdn	88.0	35.0	7.0	8.0	7.5	.81	11.0
n		Rge	25–129	11–49	1–13	5–13	2–12	.50–.94	1–26

[1]Numbers in ( ) represent per cent of total N; all others are absolute numbers.

trends found consistently in the other categories of response. The trends are seen under columns R, F+, M,FC integration[6], and S. In these areas the initial Rorschach examination by INM and the final one by RMC are most alike. Apparently the removal of the attitude of hostility was not effective, and the subjects maintained their set during the third Rorschach examination, though the suggestion had supposedly been removed following the second session. This inference is strengthened by the approximate return of category levels in condition IV with RMC to the levels established with INM initially in condition I.

It would seem, then, that (1) the subjects increased their responsiveness (R) to the stimulus cards under stress of hostility toward the examiner, (2) this change in test behavior was maintained even though the suggestion of hostility apparently had been removed, and (3) the subject's behavior did return to approximately the same level as reported initially even though a second examiner presented the stimuli. This sequence of behavior was observed also with respect to the levels of intellectual control and ego strength (F+) under stress, the amount of fantasy activity (M), the degree of emotional control as expressed in the manner in which response to form and color is integrated (FC integration), and in hostility (S) as reflected in the selection of white space in responding to the various areas of the stimulus cards.

That the mean changes seen in Table I are only trends, and not more significant changes, is reinforced by a series of analyses of variance (2) of the Rorschach responses. Taking into account the experimental design in which each of the five subjects appeared four times, therefore the necessity for correlating their successive Rorschach behaviors (in each instance it was the same individual seen four times and not four different subjects), the analyses did not reveal significant differences in the successive behaviors. That is, the variations in levels of response reported in Table I are variations which can be attributed to chance variation and not necessarily related to changed experimental conditions. The F-ratios were of the order of .50, a value much below the level of 5.32 necessary to show significant differences among the four experimental conditions.

The consistency of personality structure as reflected in the Rorschach responses, despite the experimentally changed conditions, was not unexpected. Several years ago, Leving, Grassi, and Gerson[4] reported hypnotic experimentation with induced mood changes and concluded from analysis of the Rorschach data from a series of examinations

in varying affective reactions: " . . . similarity between records is great enough to show that we were observing the same underlying personality in the various emotionally charged test situations." This observation is very much on the order of that which we experienced in the present study. Thus, the often-advanced hypothesis that the Rorschach method indicates basic personality structure[6] seems tenable for our data, i.e., although the psychiatric interview material showed changes in so-called "surface" hostility level, the underlying level characteristic of the individual did not change significantly during the several psychological examination situations.

Summary

The present paper describes a method for examining experimentally the relationship between Rorschach responses and other behavior of individuals in specific, induced conflict situations involving feelings of hostility. The technique of hypnosis was selected to alter the intensity of these feelings. By means of the induced conflict and repeated administration of the Rorschach method, the design permitted a measure of changes, if any, paralleled by psychiatric interview material.

The psychiatric interview material revealed in four of five subjects amnesia for the artificially induced conflict, and behavioral changes were definite. The fifth subject, although she did not have complete amnesia for the conflict and remembered the post-hypnotic suggestion for her to be angry with the examiner, did report tension and anxiety during the post-conflict session, in contrast with the other sessions when she had felt at ease.

The analyses of the Rorschach responses of the five subjects indicated changes in responsiveness, intellectual control, fantasy, activity, emotional control, and level of hostility. However, these changes represented trends only and there were not significant differences of response among the four experimental conditions.

The consistency of personality structure as reflected in the Rorschach records, despite the experimentally changed conditions, was not unexpected. An earlier study, by Levine, Grassi, and Gerson, has reported similar findings. Thus, the often-advanced hypothesis that the Rorschach method indicates basic personality structure seems tenable for our data, that is, although the psychiatric interview material showed changes in so-called "surface" hostility level, the underlying level characteristic of the individual did not change significantly during the several psychological examination situations.

Bibliography

1. Beck, S. J. *Rorschach's Test: II. A variety of personality pictures.* New York: Grune & Stratton, 1947.
2. Edwards, A. K. *Statistical analysis for students in psychology and education.* New York: Rinehart, 1946.
3. Erickson, M. H. The method employed to formulate a complex story for the induction of an experimental neurosis in a hypnotic subject. *J. gen. Psychol.*, 1944, 31, 67-84.
4. Levine, K. M., Grassi, J. R., & Gerson, M. J. Hypnotically induced mood changes in the verbal and graphic Rorschach, A case study. *Rorsch. Res. Exch.*, 1943, 7, 130-144.
5. Rorschach, H. *Psychodiagnostics.* Berne: Hans Huber, 1942. (Transl. by Lemkau, D., & Kronenberg, B.)
6. Williams, M. An experimental study of intellectual control under stress and associated Rorschach factors. *J. consult. Psychol.*, 1947, 11, 21-29.

A STUDY OF THE EFFECT OF COLOR ON PRODUCTIVITY AND REACTION TIME IN THE RORSCHACH TEST

RAPHAEL J. DUBROVNER, WILLIAM J. VON LACKUM, AND HUDSON JOST*

University of Tennessee

PROBLEM

Purpose. It was the purpose of this experiment to study the problem of color's influence on the individual Rorschach test by altering the stimulus cards so as to eliminate the various colors.[1] By this means some evidence could be gathered with regard to the following specific questions: (1) Does color affect the number of responses produced by the subject? (2) Does color affect the subject's time of first response?

Need for study. In the interpretation of data derived from the Rorschach test, considerable importance is attached to color factors. It is assumed that the presence of color on various cards elicits the kinds of responses that indicate the subject's habitual reaction to emotionally exciting events(1). Some aspects of emotional life are thought to be revealed by responses involving a direct use of color, such as "blood," "pink elephants," and "green caterpillars." A study of these relationships is necessarily beyond the scope of the present work.

Certain other aspects of emotional life, however, are said to be revealed by color only in an indirect way. The estimate of "affective spring," or "emotional endowment," for instance, is based upon the number of responses of any sort occurring to the all-color cards, VIII, IX, and X. Another example is the concept of "color shock," which is considered to be a neurotic pattern. Of the various possible manifestations of this shock, two major ones are fluctuations of both productivity and reaction time within the record. Color, it is stated, may increase the responsiveness of some subjects, yet diminish that of others. It may, even within a single record, show both these opposing influences. The time of first response likewise may be accelerated, retarded, or varied alternately by color(2).

An hypothesis of this nature is obviously difficult to investigate. Yet some attempt at verification would seem mandatory, for it is a tenuous assumption that the mere presence of color on several cards is responsible for certain effects peculiar to them. Other variables such as the unique shape of the figures or their brightness, may really be the crucial factors.

Literature. Most of the research on the problem of color in the Rorschach has been directed toward investigating the validity of using color as an indication of emotionality. Studies have tended to show that judgments of emotional life based on the Rorschach correspond with other clinical data; that selectivity for color in perception and in drawing may be consistent traits of an individual; that, in line with Rorschach, children prefer color to form in matching experiments less often as a function of increasing age. These studies have been summarized by Rickers-Ovsiankina(9) and by Lindberg(7), and since they do not bear directly upon the work of this paper, have been omitted here.

Of more immediate relevance is an experiment performed by Rabin and Sanderson(8). To test the belief that "the color is very carefully distributed within the ten cards in such a way that the sequence of the cards tends to enhance the color effect"(5), *p. 205*, these workers administered the Rorschach to 34 subjects with the order of cards reversed (X — I). Finding that the suddenness of the appearance of color elements in contrast to shade did not appear to affect either productivity or reaction time, they concluded that considerable doubt must be cast upon the concept of "color shock."

In a personal communication received by the author while this paper was in

*From the Psychophysiological Laboratory of Gailor Psychiatric Hospital, and the Department of Psychiatry of the University of Tennessee College of Medicine, Memphis, Tennessee. The authors wish to express their thanks to Dr. T. S. Hill, Professor of Psychiatry, and the members of the Department of Philosophy and Psychology at the University at Knoxville for their help and encouragement.

[1]Color in Rorschach usage, as defined by Beck, excludes black and its gradations.

progress, Lazarus[6] reported some of the results of a doctoral dissertation bearing on "color shock." He prepared photographic reproductions of the Rorschach cards for presentation in the Harrower-Erickson[3] group method, and compared them to the ordinary slides. His findings that color had little effect on scoring categories other than those actually containing that element (FC, CF, C) led him to suggest a revision of concepts related to color, color shock, and shock in general. The use of the group method, however, precluded an accumulation of data on the time of first response.

Method

Materials. The Rorschach stimulus figures were reproduced photographically in black and white and mounted on cards approximating the standard size. Although the reprints were reasonably exact, several artifacts could not be eliminated. As would be expected in such an achromatic reproduction, precise similarity to the originals was not obtained in certain shading effects, and in the contrast of areas within certain ink blots ordinarily due to multiple colors.

Subjects. Thirty female student and graduate nurses served as subjects for this experiment. Their ages ranged from 18 to 34, with a mean of 21 years 6 months. Most of these persons occupied the same living quarters, a circumstance which may have introduced an experimental error in that the opportunity existed for mutual discussion of the test. In order to minimize this influence, however, each subject was explicitly asked to refrain from such comment.

Procedure and controls. The subjects were divided randomly into groups A and B. The standard Rorschach cards were administered to group A, whereas group B received the facsimile presentation. An attempt was made to control the possibility of differential selection by conducting a retest in which group A received the facsimile cards and group B received the original ones. The mean time of the interval between test and retest was 14 days for both groups, the time ranging from 6 days to 32 days. The entire series of ten cards was administered in all cases in order to preserve the "temporal gestalt" of the usual clinical presentation and in order to allow a comparison of shade cards.

No changes were made in the usual instructions for the initial tests. The administration was altered only to the extent of discontinuing encouragement of the subject to respond after the first card. Instructions to the subjects for the retests were amended to inform them that the forthcoming blots were similar to those previously presented, but that former responses need not affect their present ones. Upon this occasion, no encouragement to respond was given at all.

Results and Discussion

The data presented here are grouped according to effects on productivity and effects on reaction time.

Productivity. The administration of the achromatic series of Rorschach cards does not seem to reveal important shifts in productivity (see Table 1). An examination of the table shows that on the initial test the number of responses on the "color cards" (II, III, VIII, IX, X) is greater than that on the "non-color" cards (I,IV, V, VI, VII) within group A and within group B also. On retest the results show the same tendency of the "color cards" to produce more responses than the "non-color cards" regardless of the presence or absence of color.

The comparison between the two groups in performance may be made next. For the first administration group B is somewhat superior to group A on the "color cards," and on the "non-color cards" as well. When group B's achromatic series is exchanged for the chromatic series in the retest, the superiority is still displayed. Although these differences do not appear to be significant, their precise probability of occurring through chance can not be determined statistically. Such a strict evaluation of these data would require the computation of the reliability of a difference between two differences.

Another view of the situation is given by the ranks of the cards in Table 1. Card X is consistently and unquestion-

TABLE 1. COMPARATIVE PRODUCTIVITY, BASED ON MEAN NUMBER OF RESPONSE TO CARDS, AND RANKS OF CARDS

		Non-Color Cards						*Color Cards*						*Total Mean No. Resp.*
		I	*IV*	*V*	*VI*	*VII*	*Sum of Means*	*II*	*III*	*VIII*	*IX*	*X*	*Sum of Means*	
Group A														
Chrom.	Mean	2.6	2.1	2.0	2.5	2.4	11.6	2.4	3.1	3.1	3.6	5.3	17.5	29.2
	Rank	5	4	10	6	7.5	...	7.5	3.5	3.5	2	1	...	...
Achrom.	Mean	3.3	3.5	2.7	2.9	2.9	15.3	2.6	3.9	2.9	4.3	7.3	21.0	36.4
	Rank	5	4	9	7	7	...	10	3	7	2	1	...	...
Group B														
Chrom.	Mean	2.9	3.5	2.6	2.8	2.9	14.7	2.9	3.5	3.5	2.8	6.0	18.7	33.6
	Rank	6	3	10	8.5	6	...	6	3	3	8.5	1	...	...
Achrom.	Mean	4.2	3.8	2.7	3.1	3.1	16.9	4.4	4.1	3.7	3.7	6.3	22.2	38.2
	Rank	3	5	10	8.5	8.5	...	2	4	6.5	6.5	1	...	...

ably the most productive, an observation which is in harmony with general findings. The fact that this superiority occurs during both types of administration indicates that if color does influence productivity on card X, it is not the crucial factor. Card V is apparently the least productive, another finding substantiated by Beck[1] and others. It may be noted here that card I, by virtue of its position in the series, may serve at its first appearance as a rough index of the comparability of the two groups. Table 1 shows that there is close agreement between the groups on this card, the mean number of responses differing by .3. Two noticeably large discrepancies in productivity occur, however, to cards IV and IX on the original administrations. On retest, cards II and IX show the widest disagreement. In the case of card IX the high production of responses by group A is decidedly atypical to the data reported in the Rorschach literature[1]. This may be due to the inadequacy of the present sample, but in any event, these changes do not show a consistent pattern corresponding to the changes in method of presentation.

TABLE 2. COMPARISON OF PERCENTAGES OF RESPONSES TO CARDS VIII, IX, X

	Group A	*Group B*	*Critical Ratio*
Initial Tests (chrom.-achrom.)	41.1%	36.7%	1.38
Retests (achrom.-chrom.)	39.7%	35.0%	1.64

Table 2 is of especial interest in consideration of the importance of the VIII-IX-X percentage in clinical usage. The critical ratio computed for the difference between percentages of groups A and B on the original presentation does not reach statistical significance. Whatever effect may be accredited to the influence of color, moreover, is further nullified by the retest results which reveal a difference again in favor of group A.

A further evaluation of the data covering productivity may be made by computing product-moment correlations between the number of responses on the initial tests of both groups, on the retests on both chromatic administrations, and on both achromatic administrations.

The computations show that the correlation between:

1. The chromatic series of group A and the chromatic series of group B is .69;
2. The achromatic series of group A and the achromatic series of group B is .81;
3. The chromatic series of group A and the achromatic series of group B is .72.
4. The achromatic series of group A and the chromatic series of group B is .72.

These results indicate that the correlations between sets of cards that are different with respect to the presence of color are not significantly lower than the correlations between sets of cards that are identical.

TABLE 3. MEANS OF FIRST RESPONSE TIME IN SECONDS

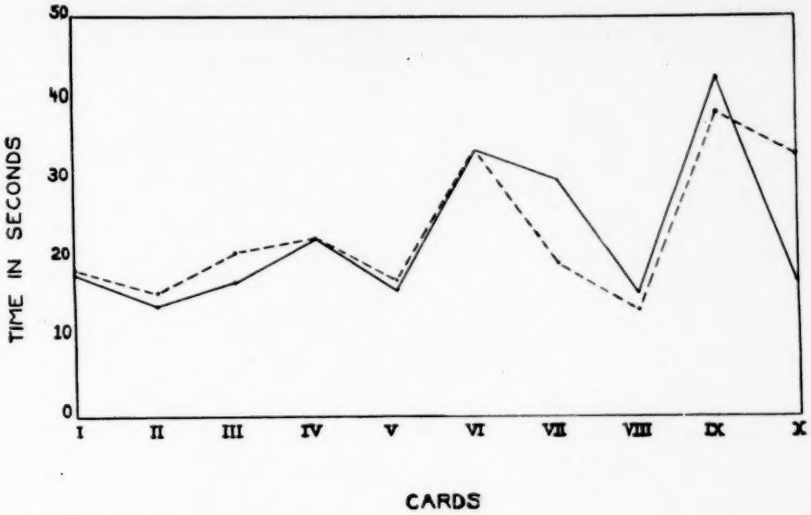
	Non-Color Cards						*Color Cards*						*Mean*
	I	*IV*	*V*	*VI*	*VII*	*Mean*	*II*	*III*	*VIII*	*IX*	*X*	*Mean*	*Total*
Group A													
Ch.	18.1	22.3	16.2	33.9	29.9	24.1	14.3	17.1	15.6	42.9	17.3	21.4	22.8
Ach.	24.1	36.0	15.3	19.1	17.3	22.4	18.8	19.7	22.3	34.2	17.0	22.4	22.4
Group B													
Ach.	18.7	22.6	17.0	33.8	19.5	22.3	16.7	20.5	13.6	38.4	33.5	24.5	23.4
Ch.	11.4	15.3	10.2	21.7	11.7	14.1	12.6	11.4	12.0	23.2	14.3	14.7	14.4

Reaction time. Data concerning the initial time of response or reaction time are summarized in Table 3 and presented graphically in Figures 1 and 2. The difference between the two groups in mean reaction time is .6 of a second for the initial testing. On retest the difference increases to eight seconds. When the "color cards" are grouped separately and compared, there is some indication that color has manifested a stimulating effect. This apparent influence is made questionable, however, by these further observations:

1. The retardation in reaction time to the "color cards" of group A on retest occurs to "non-color cards" I and IV as well. This suggests a constant practice effect instead of a color effect despite the fact that one would expect some diminution in reaction time on retest.
2. "Color cards" IX and X do not produce a pattern parallel to the pattern of the "color card" means.

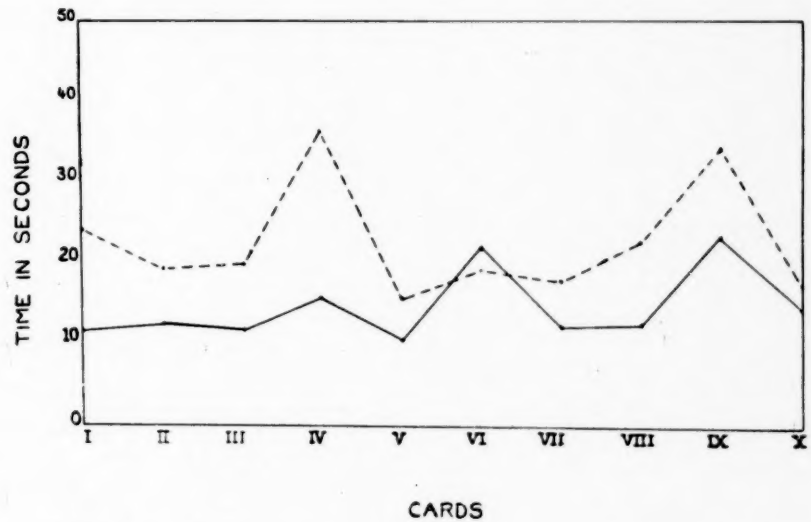
The clinical use of the contrast between average reaction time to the achromatic or "non-color cards" and the average reaction time to the cards containing color suggests a further basis for the evaluation of the present data. The contrast between the two times is assumed in clinical practice to indicate color or shade shock[10]. Upon this premise one may expect greater differences in reaction time between the two types of cards when color contrast is present than when it is not. To test this inference the computation of the sum of the differences in reaction time between "color cards" and "non-color cards" may be made without regard for signs, since the assumed capacity of color either to retard or to accelerate reaction time makes us concerned with deviations in both directions. On the original administration to both groups, results which are quite contrary to this expectation are found. The sum of the differences for group A is 105.6, whereas, group B totals 123.5. Retest reveals that again the contrast between reaction times is greater when no color is present than when both color and shade occur: group A, 69.3; group B, 58.3.

With regard to the problem of fluctuation in reaction time which is crucial to the concept of "color shock," it may be remarked that there are at least three critical points of contrast between color and shade in the sequence of standard Rorschach cards. The first point of shift is the passage from the shade card I to card II, which contains a combination of red and black elements. The second point of contrast occurs between card III, which is similar in red and black properties to card II, and shade card IV. Card VIII is the first of the final three polychromatic figures, so another shift occurs from shade card VII to card VIII. An examination of Figure 1, in which reaction times are plotted, shows that group B on the initial test parallels group A in fluctuations at these points, even though no color contrast existed. This parallel is not so clearly shown on retest (see Figure 2), but it is probably safe to suspect that practice has sufficient effect on reaction time to distort the results. It seems indicated, therefore, that certain cards are ordinarily easier or more difficult to respond to regardless of the presence or absence of color.



GROUP A: CHROMATIC SERIES - - - - - GROUP B: ACHROMATIC SERIES ———

FIG. 1. *Mean time of first response to Rorschach cards on initial test*



GROUP A: CHROMATIC SERIES - - - - - GROUP B: ACHROMATIC SERIES ———

FIG. 2. *Mean time of first response to Rorschach cards on retest*

Evaluation. It can be seen that none of the foregoing data offer support for the hypothesis that the presence of color on Rorschach figures influences either productivity of responses or speed of reaction time. The results appear to indicate that the set of cards, even in the absence of color, elicits considerable fluctuations with respect to both these factors. Some doubt, therefore, may be attached to those clinical concepts, such as "color shock" and "affective spring," which depend upon them.

An evaluation of these experimental data must not overlook several qualifying factors, however. As mentioned previously, the rationale of the Rorschach test assumes that color may have two diverse effects. It may either stimulate the individual or retard him in productivity and in response time. Color effects could thus possibly cancel themselves out in group statistics. Furthermore, interpretations of individuals' test-retest performances are distorted by the effect of repetition. Hertzman[4] has cited the practice effects listed by Harrower-Erickson[3], among which is an increase in response total. Rabin[8], too, has attributed some of his findings to this factor.

The limitations of the present technique also must be borne in mind. Seemingly minor artifacts in the photographically reproduced cards may very conceivably have exercised undue effect. Further, it is possible that the entire set of negative findings reflects only the inadequacy of the number of subjects. Finally, caution is advisable in generalizing beyond the results obtained in the present study with non-clinical subjects to the effect of color on pathological groups. Further investigations of a nature similar to this one with psychotic subjects may bear fruitful results in future research. In spite of these limitations, however, a fair judgment in view of the data is that color may not be a universal symbol. It may be an important indicator for some personalities, while having a negligible influence in determining others.

Summary and Conclusions

This experiment was designed to study the effect of color on productivity and reaction time in the Rorschach test by reproducing the standard inkblots photographically. A group of thirty nurses within an age range of 18 to 34 years was subdivided into two equal sections. The ordinary Rorschach cards were administered to group A followed by the achromatic series after an interval averaging two weeks. Group B received the achromatic presentation initially, followed after approximately two weeks by the standard series.

The results indicate that: (1) there is no evidence in the present study to support the hypothesis that color affects productivity; (2) there is no support for the interpretation of the percentage of responses to color cards VIII, IX, and X as an index of emotional resources; (3) there is no confirmation for the assumption that color affects reaction time; (4) since regular fluctuations in the difficulty of the cards occur during the achromatic presentation, it would seem that they can not be attributed solely to the effects of color, thus adding a complicating factor to the concept of "color shock."

Bibliography

1. Beck, S. J. *Rorschach's test.* New York: Grune and Stratton, 1945. Vol. II.
2. Bochner, R., and Halpern, F. *The clinical application of the Rorschach test.* New York: Grune and Stratton, 1945.
3. Harrower-Erickson, M. R. Directions for administration of the group Rorschach test. *Rorschach Res. Exch.*, 1941, 5, 145-153.
4. Hertzman, M. A. Comparison of the individual and group Rorschach tests. *Rorschach Res. Exch.*, 1942, 6, 89-108.
5. Klopfer, B., and Kelly, D. M. *The Rorschach technique.* New York: World Book Co., 1942.
6. Lazarus, R. S. An experimental analysis of the influence of color on the protocol of the Rorschach test. Unpublished Doctoral dissertation, Univ. Pittsburgh, 1948. Cited in personal communication.
7. Lindberg, B. J. *Experimental studies of colour and non-colour attitude in school children and adults.* Copenhagen: Levin and Munksgaard, 1938.
8. Rabin, A. J., and Sanderson, M. H. An experimental inquiry into some Rorschach procedures. *J. clin. Psychol.*, 1947, 3, 216-225.
9. Rickers-Ovsiankina, M. Some theoretical considerations regarding the Rorschach method. *Rorschach Res. Exch.*, 1943, 7, 41-53.

A STUDY OF THE DIAGNOSTIC POSSIBILITIES OF THE CVS WITH A GROUP OF ORGANIC CASES[1]

JOSEPH D. MATARAZZO

Northwestern University

Introduction

With the development of intelligence scales such as the Stanford-Binet and Wechsler-Bellevue, there have been numerous studies on the significance of "scattering" on these scales. However, until recently the results of these scatter studies have been both confusing and contradictory[4, 5, 6, 17]. While this might be due, as Jastak[13] points out, to the failure of the experimenters to pay attention to certain technical problems, the most probable reason is that the tests used are not suitable for scatter studies. One of the chief difficulties has been the fact that with scales such as the Stanford-Binet, the measurement of intelligence was of primary interest and the diagnostic possibilities of scatter were only of incidental interest. The omnibus-type test yields a valid total score but confuses the identity of various specific abilities by not isolating them into subtests. It is thus only natural that scatter studies employing this test have yielded differing or negative results[4, 5].

The introduction of the Wechsler-Bellevue scale[17] partly overcame the difficulties for scatter analysis presented by the omnibus-type scale. However, scatter studies employing the Wechsler-Bellevue scale have yielded somewhat discouraging results, in part because of the complexity of possible interrelationships afforded by the combination of eleven subtests[3, 15, 16].

The recent development of abbreviated tests have shown that it is both possible and practical to devise batteries consisting of a smaller number of subtests. If these subtests are selected not only for their potentiality in measuring intelligence, but also for their clinical potentiality as differential clinical indicators[12, 13] the possibility of their providing valid scatter measures should be enhanced.

This double principle of selection has been used in the development of a series of such abbreviated scales at Northwestern University[7, 8, 9, 10, 11]. A promising one, the CVS, consists of the Comprehension and Similarities subtests of the Wechsler-Bellevue scale[17], and a Vocabulary test made up of a fifteen-word list proposed by Thorndike. As stated elsewhere[9, 10, 11], this battery of three subtests was chosen for its correlation with other intelligence tests (+.76 with both the full Wechsler-Bellevue scale and the Navy GCT test), as well as for its diagnostic potentiality, since it contains a vocabulary subtest which classically has been found to be relatively insensitive to psychopathosis [1, 9, 10, 16] and comprehension and similarities subtests, both of which have been found to be sensitive to psychopathosis[9, 10, 15, 16, 17]. The CVS, therefore, is clinically useful since it affords a "scatter" measure which is defined as the discrepancy between performance level on Vocabulary and that on the Comprehension and Similarities tests.

Its use as a diagnostic instrument, through the interpretation of scatter scores, has thus far been reported on only two deviant groups—a schizophrenic and an aged group[9, 10]. The scatter measure used in these previous studies was the presence of C and S scores both below the Vocabulary score. Since past results with the CVS battery have been very encouraging, it is desirable to extend our knowledge of the diagnostic possibilities of the test to other clinical groups, particularly organic cases. The present study concerns itself with the investigation of the possibilities of the CVS scatter score in detecting organic pathosis in a paretic group.

[1]This study is part of a larger project under the supervision of Dr. William A. Hunt and subsidized by the Office of Naval Research under their policy of encouraging basic research. The opinions expressed, however, are those of the individual author and do not represent the opinions or policy of the Naval service.

Subjects

The experimental group consisted of 50 hospitalized paretic patients.[2] There was unanimous staff agreement on this diagnosis for each of the patients. The ages ranged from 33 to 55 with a mean age of 48.38 and an S.D. of 6.02. The distribution was skewed negatively as is evidenced by a median age of 50.0 for the group. The educational level for this group ranged from 0 to 14 years with a mean of 8.68 years completed and an S.D. of 3.06. The distribution of educational level was essentially normal. The mean number of years of hospitalization for these organics was 9.68 with an S.D. of 5.46 and a median falling in the 9 to 10 range. Included in this paretic group were only those patients who were of the white race, born in the U. S., and who were, at the time of testing, cooperative enough to yield valid test results.

The control group consisted of 50 non-institutionalized "normals" composed of a sample of the firemen from three different cities.[3] Nineteen of these firemen were from the Schenectady Fire Department, 12 from the Evanston Fire Department, and 19 from the Chicago Fire Department. Included in this "normal" group were only those firemen who were of the white race, born in the U. S., had no previous history of organic brain damage, and who at the time of testing showed no obvious evidence of organic brain pathosis. The range of ages for the firemen, like the organics, was from 33 to 55, with a mean age of 46.32, an S.D. of 5.26, and a median of 46 to 47. The mean number of years of school completed was 9.06 with an S.D. of 2.02 and a median of 8 to 9 for this group.

Fisher's t was applied to the mean ages and educational levels of the two groups and the values were found to be 1.81 and 0.73 for the two variables respectively. Neither of these two variables is significant at the 5% level. The variances of the two distributions of ages were compared by the t-test and the differences were found not to be significant at the 5% level. Thus we may conclude that our two groups were matched for age. As regards educational level, however, our two groups were not so adequately matched. Although the mean educational levels of both groups were found not to be significantly different, a t-test of the sigmas of both groups showed that these distributions *did* differ significantly in regard to their variances. Thus we must conclude that our two groups were not adequately matched for educational level and, by inference, that they were not equated for intelligence. As will be seen later, this latter fact has little bearing on the present study.

In conclusion we may say that our experimental and control groups were adequately matched for sex, race, and age and somewhat less adequately matched for educational achievement, and thereby meet the standards set by Jastak for experimentation on test scatter[13].

Procedure

The CVS was administered individually to each of the 50 paretics and the 50 firemen. The order of presentation of the CVS was alternated for each successive subject. All the subjects were tested singly in a separate room by the experimenter and the standard testing and scoring procedures for the CVS were used[8].

Results and Discussion

Table I presents the mean scores and standard deviations for C, V, S and CVS (total CVS score) for both the paretic and control groups. Also included in this table are the Fisher t-ratios for the significance of the differences between the mean values of the two groups.

Inspection of this table shows that the paretics made mean scores of 4.84, 8.86, and 3.66 on C, V, and S, respectively, while the controls made scores of 11.62, 12.88, and 10.70 on the same three subtests. The differences of means for these three tests are all significant (1% level)

[2]Thanks are due to Dr. S. G. Klebanoff for his cooperation and assistance and to Dr. Phyllis Wittman and the staff of the Elgin State Hospital for the use of their facilities.
[3]Thanks are due to the officers and men of the Schenectady, Evanston, and Chicago Fire Departments for their interest and cooperation in this phase of the study.

TABLE I

MEANS, SIGMAS, AND T – RATIOS FOR THE PARETIC AND CONTROL GROUPS ON THE CVS

	Paretics		*Controls*		*t-ratio*
	Mean	Sigma	Mean	Sigma	
C	4.84	2.65	11.62	1.97	14.42*
V	8.86	3.57	12.88	2.41	6.48*
S	3.66	2.52	10.70	2.56	13.54*
CVS	17.36	7.36	35.20	5.58	13.52*

* Difference between two means is significant at the 1% level of confidence.

as can be seen from the column of *t*-ratios.

Before we assume that the differences within the paretic group indicate the the presence of scatter, we should consider the 4.02 difference between the paretics' Vocabulary score and that of the controls. Since Vocabulary is known to correlate quite highly with intelligence, and since Vocabulary is relatively insensitive to deterioration, this difference between the Vocabulary scores of the two groups indicates that our two groups are not equated for intelligence—the controls apparently being more intelligent. This finding is not surprising since firemen are probably above the average in intelligence for the general population in their age range. However, the fact that the controls appear more intelligent in no way affects the validity of the present study, since the hypothesis being tested is that there will be greater scatter or spread of scores among the paretics than among the controls. This means that we are interested not in comparing the mean scores of each group, but in comparing the mean difference between Vocabulary and other scores (C and S) in each of the two groups.

Examination of Table II shows that for the paretics the difference in means

TABLE II

DIFFERENCES IN MEANS OF C AND V, S AND V, AND C AND S: THE SIGNIFICANCE OF THESE DIFFERENCES: AND THE SIGNIFICANCE OF THE DIFFERENCES BETWEEN THESE DIFFERENCES FOR PARETICS AND CONTROLS

	C and V				
	Mean C	Mean V	V-C Diff	σDiff	*t*-ratio
Paretics	4.84	8.86	4.02	0.60	6.70*
Controls	11.62	12.88	1.26	0.44	2.86*
Diff			2.76		
σDiff			0.74		
t-ratio			3.73*+		
	S and V				
	Mean S	Mean V	V–S Diff	σDiff	*t*-ratio
Paretics	3.66	8.86	5.20	0.59	8.81*
Controls	10.70	12.88	2.18	0.49	4.45*
Diff			3.02		
σDiff			0.77		
t-ratio			3.92*+		
	C and S				
	Mean C	Mean S	C–S Diff	σDiff	*t*-ratio
Paretics	4.84	3.66	1.18	0.50	2.36**
Controls	11.62	10.70	0.92	0.46	2.00**
Diff			0.26		
σDiff			0.68		
t-ratio			0.38+		

* Difference between two means is significant at the 1% level of confidence.

**Difference between two means is significant at the 5% level of confidence.

+The two populations of differences were found to be of equal variance since a *t*-test found the sigmas not to be different.

between the Vocabulary and Comprehension scores is 4.02; i.e., the paretics made a mean V score 4.02 units higher than on C. This difference is seen to be significant at the 1% level. The control group showed a similar scatter as can be seen by their V-C difference which is 1.26. This difference of 1.26, though itself significant at the 1% level, is still 2.76 units lower than the same V-C difference for the paretics. As is shown in the vertical difference column of Table II, this difference of 2.76 between the V-C differences of the paretic and control groups is significant at the 1% level. While the normal control group thus does show a statistically significant difference between V and C as measured by our test (other studies have shown a similar normal difference between V and C with advancing chronological age) this difference is not as great as that exhibited by the paretic group where the effects of organic pathosis are added to the effects of normal aging.

Similar results are seen when the mean S score is compared to the mean V score in both groups and this difference between V and S in one group is compared to the V-S difference for the other group. Here again the paretics show a greater scatter than do the normals.

Caution is necessary in assuming that these absolute score differences can be arbitrarily equated with the clinical significance of the organic deterioration assumed to be present, since, as Jastak[13] has shown, absolute score difference are not of equal significance at different ends of the distribution. Thus, a difference between a V score of 18 and a C score of 16 would be mathematically equal to the difference between a V score of 4 and a C score of 2 but one would expect a more significant clinical change in behavior in the latter case. Later on, in using this scatter diagnostically, we will make use of a formula to correct for this.

Finally, Table II also answers the question as to whether or not the deficit shown on C is greater than that shown on S. The deficit in performance for both groups is not significantly greater for C than it is for S, indicating that both tests are showing about the same amount of deficit in the paretic and control groups.

While the differences mentioned above bear out our hypothesis that paretics show greater "scatter" on the CVS battery than does a control group of the same age, we have thus far only demonstrated this as a group measure. It thus becomes necessary to make a prediction, not concerning group differences, but concerning a single individual. It is necessary to introduce some evaluation of the performance of the test in predicting psychopathosis in an individual patient. The standard technique in current use for doing this is the application of some criterion to the test performances of both a clinical group and a normal group and the determination of the number of clinical cases which are detected by the test as well as the number of "false positives."

In a previous investigation of the diagnostic use of the CVS battery with psychotic cases[10], the criterion for psychopathosis used was the appearance of both C and S scores below V. Table III shows the results of applying this measure to our groups. While the percentage of paretics detected by the measure (84%) can be considered reasonably efficient for a brief test of this sort, the percentage of false positives is too high (48%). This result is to be expected since it could be predicted (and Table I demonstrates) that the normal group, due to its advanced age, would also show scatter. It is obvious that we need a differential measure which will recognize not the mere existence of scatter, but the greater amount of it present in the paretic as compared to the control

TABLE III
NUMBER AND PERCENT OF CASES HAVING C AND S SCORES BOTH LESS THAN V

	N	%
Paretics	42	84
Controls	24	48

group, a measure which will give practical diagnostic significance to the group differences reported in Table I.

The measure, C+S – V, which was proposed by Hunt and French in their investigation of the diagnostic possibilities of the CVS with a functional group[10], offers us a possible lead. This scatter measure was designed on an *a priori* basis; i.e., since C and S classically have been shown to deteriorate with pathosis while V has been shown to be less sensitive to deterioration, this measure (a weighted C and S score from which is subtracted the individual's V score) will yield a satisfactory index of scatter. Ideally an individual should make a similar score on each of the three subtests resulting in a scatter score of O.O. A score below O.O (a negative score) is defined as scatter. At present the significance of a positive score resulting from this measure is unknown. Table IV shows the mean amount of scatter, $\frac{C+S}{2}$ – v, for the paretic and control group when this measure is applied to each individual case and the resulting individual amounts of scatter are then averaged in order to obtain a measure of the group trends. This measure also reveals the greater scatter for the paretic group as compared to the control group; i.e., the paretics showed a mean amount of scatter of –4.59 while the controls showed a mean of only –1.82, the difference being significant at the 1% level. However, as we stated above, we are interested in the diagnostic potentiality of our measure for *individual* predictions. A method which will enable us to interpret group trends for individual predictions is the use of a cut-off point. The cut-off point selected was a $\frac{C+S}{2}$ – v score of –3.77 which was one standard deviation (1.95) below the control group's mean. The individual data were then examined and as Table V shows, 62% of the paretics showed scatter scores greater than –3.77 while only 12% of the controls showed this great a scatter. These results, a pick-up rate of 62% and a false positive rate of 12%, are within the limits which have previously been accepted as serviceable and further substantiate the diagnostic practicability of the abbreviated battery, CVS. It would be desirable, however, to increase the pick-up rate of 62% if possible.

TABLE V.

NUMBER AND PERCENT OF CASES HAVING $\frac{C+S}{2}$ – V SCORES LESS THAN –3.77*

	N	%
Paretics	31	62
Controls	6	12

*This cut-off point is one sigma (1.95) below the controls' mean of –1.82.

Another possible measure would be a scatter ratio such as that previously mentioned by Jastak which corrects for the fact that absolute score differences are not of equal significance at different ends of the distribution. We used the ratio $\frac{C+S}{2} \div v$ since in cases of no deterioration, i.e., young normals, the three subtest scores will ideally be equal, yielding a ratio of unity (1.0). In cases of deterioration we should expect, since C and S are known to deteriorate while V holds up, a scatter ratio below 1.0.

TABLE IV

TABLE SHOWING MEAN AMOUNTS OF SCATTER AND MEAN SCATTER RATIOS FOR PARETIC AND CONTROL GROUPS

	Paretics		*Controls*		*t-ratio*
	Mean	Sigma	Mean	Sigma	
$\frac{C+S}{2}$ – V	–4.59	2.87	–1.82	1.95	5.54*
$\frac{C+S}{2} \div V$	0.51	0.24	0.88	0.15	9.52*

*Difference between two means is significant at 1% level.

Applying this ratio to our group data yielded the results shown in Table IV. As was the case with the scatter measure, $\frac{C+S}{2} - V$, the paretics showed more mean scatter than did the controls. The difference between the paretics' mean scatter ratio of 0.51 and that of the controls' 0.88 is seen to be significant at the 1% level. The use of a cut-off point similar to the one mentioned above, i.e., one standard deviation (0.15) below the mean scatter ratio of the controls, yields the diagnostic results which are shown in Table VI. Thus we see that the scatter ratio, $\frac{C+S}{2} \div V$, picks up 82% of the paretic group and falsely identifies (false positives) 10% of the controls. The difference between the controls' mean of 0.88 and 1.0 (no deterioration) is significant[4] and we can attribute this difference to the normal effects of aging. The difference between the mean of 0.51 of the paretic group and 1.0 is significant and can be attributed to the combined effects of aging and pathosis.

TABLE VI

NUMBER AND PERCENT OF CASES HAVING $\frac{C+S}{2} \div V$ SCORES LESS THAN 0.73*

	N	%
Paretics	41	82
Controls	5	10

In order to investigate how much more successful was the scatter ratio, $\frac{C+S}{2} \div V$, in this study than was the reviously used scatter measure, $\frac{C+S}{2} - V$, a Chi square test was done to test the significance of the difference in the pick-up performance of the two tests. This difference was found to be significant at slightly better than the 5% level and we conclude that, for this study, the scatter ratio is a better diagnostic measure. Its further use with other groups will be necessary before any definite conclusions can be drawn. It should be mentioned however, that both the scatter ratio and the scatter measure of Hunt and French are diagnostically more sensitive to pathosis in older individuals than is the measure C and S both lower than V.

[4]Since no distribution is available for the hypothetical mean of 1.0 we must assume one. The standard deviation for this hypothetical distribution could be as high as .28 and still yield a difference significant at the 1% level. Such a high variability is extremely unlikely.

CONCLUSIONS

As a result of the above findings we can conclude that CVS will perform as a diagnostic indicator for screening paretics, thus confirming with organic cases the previous findings with functional cases. We must stress again, however, that in using the test diagnostically with older groups it is necessary to take account of the normal deterioration that takes place as a function of aging. In cases involving late adolescence and early adulthood the previous measure, C and S both lower than V, is to be recommended since this obviates the need of any further arithmetical treatment. This criterion is unsatisfactory for diagnostic purposes with older age groups, however, and a further refinement is necessary. On the basis of this study it is suggested that the scatter ratio, $\frac{C+S}{2} \div V$, plus a cut-off point offers a suitable means for overcoming this difficulty. Our results further confirm the desirability of age norms for scatter as was previously suggested by Jastak, and Hunt and French. As was brought out in our introduction, failure to control the age factor has been a serious handicap in many of the past studies of test performance in clinical groups.

BIBLIOGRAPHY

1. BABCOCK, H. An experiment in the measurement of mental deterioration. *Arch. Psychol. N. Y.*, 1930, 18, No. 117.
2. FOSTER, A. Age and the Wechsler-Bellevue scattergraph. *J. clin. Psychol.*, 1947, 3, 396-397.
3. GILLILAND, A. R., WITTMAN, P., & GOLDMAN, M. Patterns and scatter of mental abilities in various psychoses. *J. gen. Psychol.*, 1943, 29, 251-260.
4. HARRIS, A. J. & SHAKOW, D. The clinical significance of numerical measures of scatter on the Stanford-Binet. *Psychol. Bull.*, 1937, 34, 134-150.
5. HARRIS, A. J. & SHAKOW, D. Scatter on the Stanford-Binet in schizophrenic, normal, and delinquent adults. *J. abn. soc. Psychol.*, 1938, 33, 100-111.
6. HUNT, J. McV. & COFER, C. N. "Psychoolocal Deficit". In *Personality and the Behavior Disorders* by Hunt, J. McV., 1944, pp. 971-989.

7. HUNT, W. A. & FRENCH, ELIZABETH G. Some abbreviated individual intelligence scales containing non-verbal items. *J. consult. Psychol.*, 1949, 13, 119-123.
8. HUNT, W. A. & FRENCH, ELIZABETH G. A second fifteen-word vocabulary test for use with abbreviated intelligence scales. *J. consult. Psychol.*, 1949, 13, 124-126.
9. HUNT, W. A., FRENCH, Elizabeth G., KLEBANOFF, S. G., MENSH, I. N., & WILLIAMS, M. Further standardization of the CVS individual intelligence scale. *J. consult. Psychol.*, 1948, 12, 355-359.
10. HUNT, W. A., FRENCH, ELIZABETH G., KLEBANOFF, S. G., MENSH, I. N., & WILLIAMS, M. The clinical possibilities of an abbreviated individual intelligence test. *J. consult. Psychol.*, 1948, 12, 171-173.
11. HUNT, W. A., KLEBANOFF, S. G., MENSH, I. N., & WILLIAMS, M. The validity of some abbreviated individual intelligence scales. *J. consult. Psychol.*, 1948, 12, 48-52.
12. HUNT, W. A. & OLDER, H. J., Psychometric scatter pattern as a diagnostic aid. *J. abn. soc. Psychol.*, 1944, 39, 118-123.
13. JASTAK, J. Problem of psychometric scatter analysis. *Psychol. Bull.*, 1949, 46, 177-197.
14. KLEBANOFF, S. G. The Rorschach test in an analysis of personality changes in general paresis. *J. Personal.*, 1949, 17, 261-272.
15. MAGARET, A. Parallels in the behavior of schizophrenics, paretics, and pre-senile non-psychotics. *J. abn. soc. Psychol.*, 1942, 37, 511-528.
16. RAPAPORT, D., et al. *Diagnostic psychological testing.* Chicago: Year Book Publishers, 1945, Vol. I.
17. WECHSLER, D. *The measurement of adult intelligence.* Baltimore: Williams and Wilkins, 1944.

VALIDATION STUDIES OF THE RORSCHACH TEST: SHYNESS AND GREGARIOUSNESS IN THE NORMAL SUPERIOR ADULT[1]

WAYNE H. HOLTZMAN

University of Texas

INTRODUCTION

The present report concerns one phase of a series of studies designed to investigate the relationship between Rorschach test patterns and certain aspects of the personality of the normal, superior adult as judged by his intimate associates in everyday life. After a careful search of the literature concerning Rorschach test interpretation, ten traits were selected which could be easily judged by intelligent subjects. In addition to this, information on three other variables was collected: leadership ability, degree of familiarity between individuals, and scholastic grade point averages.

[1]Part of a dissertation presented to the Faculty of the Graduate School of Stanford University in candidacy for the Degree of Doctor of Philosophy. The author wishes to express his thanks to Professors Frances G. Orr, Katharine P. Bradway, and Ernest R. Hilgard of the Department of Psychology, Stanford University, for their helpful suggestions throughout the design and execution of this investigation. This paper was read at the meeting of the Midwestern Psychological Association, Detroit, Michigan, May 5, 1950.

METHOD

The personality judgments used as criteria for the Rorschach were obtained by having each individual within two independent groups of Stanford University men rank every member within his group, including himself, on each of the ten traits and on Leadership and Familiarity. Having an individual rank himself as well as the rest of the group apparently resulted in greater ego-involvement of the subject doing the ranking and made possible the study of discrepancies between self and group judgments of each individual. The trait descriptions defining each trait are given in Figure 1 in the order of presentation to the subject.

In order to assure reasonably adequate criterion data the two groups of college students were carefully selected from men living in Stanford Village, a housing project of Stanford University, so that each member of a group was well acquainted with all the others in his group and had lived with the group a minimum of four months. Placement of

FIGURE 1

Trait Name	*Trait Description*
Drive	Drives himself onward, always tries hard, very ambitious
Shyness	Shy, bashful, rarely participates in social functions, overcautious in making emotional contacts with other people
Compulsiveness	Somewhat compulsive, habit-bound, overly tidy, a precisionist, punctilious, fussy
Stubbornness	Stubborn, arrogant, headstrong, mulish
Immaturity	Emotionally immature, infantile, demanding, self-centered
Creativity	Capable of highly original ideas, inventive, resourceful, ingenious
Tactlessness	Tactless, insensitive to the feelings of other people, a blunderbuss, like a bull in a china shop
Gregariousness	Very gregarious, the "life-of-the-party" type, mixes well and seems to enjoy being around people.
Impulsiveness	Very impulsive, flies off the handle, gets mad easily, has poor control of emotions, temperamental.
Adjustment	Well adjusted, handles personal problems well, rarely shows signs of excessive worry or anxiety, seems basically happy
Leadership	Place the individual whom you think would be most desirable as a leader to represent the group in campus affairs at the *top*. Place the next to the best just below him and continue in this way till the entire group has been ranked.
Familiarity	Place the individual at the top whom you feel you know better than any of the others. Rank the remainder of the group below him placing the man you feel you know the least at the bottom.

students within living groups was done randomly by the University at the beginning of the year from a list of students who had applied for housing. The two groups of students which met the above requirements consisted of 24 men in Building A and 22 men in Building B.

The arrangement of living quarters at Stanford Village is ideal for a study of this type. The majority of students assigned to a wing of a building, study and sleep in one large room with only desks and lockers to separate one cubicle of students from another. Consequently the pooled rankings on such social traits as Gregariousness and Shyness can be expected to have fairly high validity.

Pooled rank scores for each individual on every trait were computed by summing the individual ranks assigned him by his associates and placing the subjects in rank order according to their pooled scores. This final rank order was the one used as the criterion for that trait. Reliability coefficients for the final pooled scores, computed by the split-half method were in general highly satisfactory for criterion purposes. Gregariousness had a reliability of .97 in Group A and .94 in Group B. The reliability for Shyness was .96 and .78 in Groups A and B respectively. Since these two traits are socially observable by their very nature and can be easily observed by the judges, the final pooled ranks for Shyness and Gregariousness should be unusually good criteria for comparison with the Rorschach.

An individual Rorschach was administered to each subject and scored with methods prescribed by Klopfer and Kelley[5]. The scoring of six records selected at random was checked independently by two other psychologists and was found to agree essentially with the investigator's scoring.

The Rorschach data were given to two psychologists expert in Rorschach interpretation, one working with each group, and each psychologist was given instructions to rank the members of his group on each of 11 traits using only the information available in the Rorschach test. The psychologists' evaluations were then compared with the criterion data. Unfortunately, the task of carefully studying so many Rorschach records and of keeping all of them in mind while attempting to rank them on each trait proved most difficult, and a more superficial analysis of each record was resorted to by the two Rorschach experts. None of the correlations between these approximate rank orders and the

criterion data for *any* of the traits were significantly different from zero, and this method of evaluation was reluctantly abandoned for the present.

A second approach to the validation of the Rorschach consisted of testing specific hypotheses regarding the prediction of each trait by objective Rorschach scores. *First*, various hypotheses as stated in published papers and books regarding the traits studied were put to test using the ranked data of Group B as the criteria. *Second*, the relationships between the Rorschach and rank data for Group B were analyzed by the experimenter *without any knowledge of the test data for Group A*, and various new hypotheses were set up to account for the relationships found between certain Rorschach characteristics and the traits. And *third*, these new hypotheses together with those tested previously on Group B were tested by the data for Group A.

Results

The two traits studied in this paper, Shyness and Gregariousness, will be considered simultaneously since the correlations between these two traits, −.92 and −.85 in A and B respectively, were nearly as high as their estimated reliabilities. All the trait intercorrelations in both groups were analyzed for clusters by Tryon's method[9]. The high congruence of the correlation profile graphs for Gregariousness and Shyness indicated that in this study Gregariousness and Shyness were essentially opposite extremes of a bipolar trait.

Previously Formulated Hypotheses. Most Rorschach workers accept the proposal made by Rorschach that the movement responses are a direct result of "promptings from within," of the workings of the imagination[1, 5, 6, 7]. Likewise, reactions to color are thought to be motivated by a responsiveness of the individual to external stimuli. From this, Rorschach deduced that the "Erlebnistyp" of the individual was related to the ration of M to sum C scores on the Rorschach test, a person with greater M than sum C tending to be an introvert and one with sum C greater than M tending to be an extravert. The first hypothesis to be tested deals with this and other purported measures of so-called "Erlebnistyp."

Hypothesis 1: *The degree of Gregariousness manifested by an individual is directly related to his Erlebnistyp; the more extraversive he is, the more gregarious he will appear to others.*

As one test of this hypothesis we have the following:

1a. *The ratio M:sum C is inversely related to the degree of Gregariousness.*

Klopfer and Kelley have extended the logic used by Rorschach to include a second measure of Erlebnistyp: "The ratio (FM+m) :(Fc+c+C') represents introversial and extraversial tendencies not fully accepted or utilized by the subject[5], *p. 253.*"

1b. *The ratio (FM+m):(Fc+c+C') is inversely related to the degree of Gregariosuness.*

A third measure of responsiveness to "stimuli from without" proposed by Klopfer and Kelley is the percentage of responses to Cards 8, 9, and 10[5], *p. 254*).

1c. *The percentage of responses to the last three cards is directly related to the degree of Gregariousness.*

A fourth variation of Hypothesis One which was tested on both groups concerned a ratio of the total number of movement responses to the total number of responses to color.

1d. *The ratio (total movement) : (total color) is inversely related to the degree of Gregariousness.*

The rank-difference correlations between Gregariousness and the four ratios tested under Hypothesis One are presented in Table 1. None of the correlations differ significantly from zero indicating that there is little, if any, relationship in this study between the indices of "Erlebnistyp" on the Rorschach test and the trait of Gregariousness or its opposite, Shyness.

TABLE I

THE RELATIONSHIP OF ERLEBNISTYP TO GREGARIOUSNESS

Hypothesis	Test Score	*Correlation with Gregariousness in* Group A (N = 24)	Group B (N = 22)
1a	M: sum C	.04	−.11
1b	(FM+M):(Fc+c+C′)	.19	.35
1c	%R to cards 8, 9, 10	.21	.24
1d	Movement: Color	.02	.15

HYPOTHESIS 2: *An individual will be judged shy if the ratio* $(Fc+c+C') : (FC+CF+C)$ *is greater than two.*

This hypothesis was derived directly from a statement by Klopfer and Kelley, ". . . the point of real 'contact shyness' or overcautiousness in emotional contacts (is reached) if the total number of achromatic responses is at least twice the total number of bright color responses[5], *p. 230.*" Since the descriptive phrase "overcautiousness in emotional contacts" was used as part of the trait description for Shyness, the criterion should be particularly appropriate for testing this hypothesis.

The above ratio was computed for each individual in both groups. Five individuals in Group B and four in Group A were found to have ratios which were greater than two. Inspection of the ranks assigned these individuals on the trait, Shyness, revealed that in Group A individual C was ranked 10th from the top on Shyness, L was ranked 12th, R was ranked 16th, and W was ranked 21st. In group B, individuals R, I, B, M, and P were ranked 3, 9, 12, 15, and 21 respectively. Since there were only 24 men in A and 22 men in B, the hypothesis was clearly unsubstantiated in both groups.

HYPOTHESIS 3: *The absence of color responses in normals is a sign of shyness.*

This third hypothesis was taken directly from Rapaport[6], *p. 248*). Although no individuals in Group A had Rorschach records with no color responses, three members of Group B used no color whatsoever. The ranks of these three subjects on Shyness were 3, 9, and 15. Since the rank of 22 is the least shy in the group, it is apparent that the absence of color in normal records is not a very useful sign of Shyness.

Empirically Derived Hypotheses. In the hope of finding some patterning of Rorschach scores or signs which would be related to Gregariousness or Shyness, the Rorschach records for Group B were analyzed carefully prior to any analysis of the records for Building A. In this way new hypotheses might be proposed which could be tested on the Rorschach data for Group A. All efforts to find an objective index of Gregariousness failed. In addition to various combinations of the hypotheses tested above, the following variables were investigated: average time per response; the ratio of average reaction time to the achromatic cards and average reaction time to the colored cards; F per cent; the ratio (H+A) : (Hd+Ad); number and quality of M; number of H; and number of P.

From this analysis of the data for Group B only one tentative hypothesis of rather restricted use was found to be feasible.

HYPOTHESIS 4: *Unless at least one human concept (whole or part of the human figure) is seen in the Rorschach test, the subject is not gregarious.*

Only one member of Group B failed to see at least one human concept. He was ranked 18th from the top on Gregariousness. Unfortunately when the hypothesis was tested on Group A it was discovered that everyone in this group had seen at least one human concept. Consequently, the hypothesis remains unverified.

Discussion

In any validation study, the nature of the criterion selected for testing hypotheses is of utmost importance. To the writer's knowledge the present investigation is the first one using pooled rank order data as criteria for the study of the Rorschach test. The method of procuring these pooled judgments of personality characteristics, however, is not unique. Several research workers have used techniques somewhat similar to the present one with varying degrees of success for the study of personality [2, 3, 4, 8]. In some respects evaluations of an individual's traits by his intimate associates are the best possible estimates of those traits. This is particularly true whenever the personality characteristic being studied is a phenotypical, socially defined trait such as Shyness or Gregariousness.

Because of the nature of the criterion, the present results strongly indicate the need for considerable caution in evaluating Shyness or Gregariousness by the Rorschach test. Some may argue that the trait ranges are highly restricted since these are college students. The high trait reliabilities provide evidence that the range was at least sufficiently wide to permit relatively precise ordering of the individuals on Gregariousness and Shyness. Be this as it may, the argument of range restriction is nevertheless irrelevant to the purposes of this study. The individual Rorschach records demonstrated extreme fluctuations on most objective scores. For example, the ratio of M to sum C scores ranged from zero to infinity with individuals having as many as 18 human movement responses or as high as 21 color responses. It is the range of the test scores which must be heterogeneous in a study of predictive efficiency.

However, this does not mean that the concept of Erlebnistyp as such is not a useful construct for personality analysis. One might have a strong tendency to rely upon fantasy for self esteem and yet outwardly appear quite gregarious. Thus a heavy predominance of movement over color might indicate the possession of a rich fantasy life to which the individual turns for maintenance of self esteem. But this same person might be quite active in social affairs and appear to his associates as anything but shy. More research is needed regarding the personality characteristics which prompt the projection of movement into responses before the concept of Erlebnistyp can be accurately evaluated.

Summary

The relationships between the Rorschach test and the traits, Shyness and Gregariousness were studied in two groups of college men. The pooled judgments of Shyness and Gregariousness proved highly adequate as criteria for testing various hypotheses relating to the prediction of these two traits by the Rorschach.

Two methods of validation were used: (1) intuitive, "blind" analyses by two Rorschach experts; and (2) quantitative and qualitative analyses of objective Rorschach signs. All attempts to find significant relationships between the Rorschach and Gregariousness or Shyness failed in spite of the criterion adequacy.

The conclusion is made that the commonly claimed relationship between Rorschach test data and the personality traits, Shyness and Gregariousness, is not supported in this study of normal superior adults.

Bibliography

1. Beck, S. J. *Rorschach's test: II. A variety of personality pictures.* New York: Grune & Stratton, 1945.
2. Cattell, R. B. *Description and measurement of personality.* New York: World Book Co., 1946.
3. Fiske, D. W. Consistency of the factorial structures of personality ratings from different sources. *J. abnorm. soc. Psychol.*, 1949, 44, 329-344.
4. Green, G. H. Insight and group adjustment. *J. abnorm. soc. Psychol.*, 1948, 43, 49-61.
5. Klopfer, B. and Kelley, D. *The Rorschach technique.* Yonkers, New York: World Book Co., 1942.
6. Rapaport, D. *Diagnostic psychological testing: volume II.* Chicago: The Yearbook Publishers, 1946.
7. Rorschach, H. *Psychodiagnostik: Methodik und Ergebnisse eines Wahrnemungsdiagnostischen Experiments.* Berne, Switzerland: Huber, 1941 (Fourth Edition). English edition, Berne, Switzerland: Huber, 1942.
8. Sears, R. *Experimental studies of projection: I. attribution of traits.* J. soc. Psychol., 1936, 7, 151-153.
9. Tryon, R. C. *Cluster analysis.* Ann Arbor: Edwards, 1939.

VALIDATION STUDIES OF THE RORSCHACH TEST: IMPULSIVENESS IN THE NORMAL SUPERIOR ADULT

WAYNE H. HOLTZMAN

The University of Texas

Introduction

Among the numerous studies concerned with the Rorschach Test those dealing with the reaction of individuals to color in the ink blots have stimulated considerable controversy. Rorschach assumed that stimulus-color often played a critical role as a determinant of responses to ink blots. Furthermore he recognized other effects of the color which he grouped together under the term "color shock." Klopfer and Kelley[2] discuss the significance of color responses at considerable length, proposing that bright color reactions are indicative of emotional responsiveness to the environment. Color responses in which form apparently plays a minor role are said to be indicative of a more impulsive emotional reaction than those having definite form characteristics. In a recently reported experiment[3], Lazarus attempted to test the influence of color on various scoring categories by presenting two sets of slides, one consisting of the standard Rorschach ink blots and the other consisting of achromatic reproductions of the originals, to two matched groups of high school students. He found no significant differences between the colored and achromatic sets with regard to the usual indices of "color shock." It should be pointed out, however, that the more significant indicators of color disturbance, such as emotional exclamations referring to color, delayed reaction times to colored cards, and other such behavior obtainable only when the test is administered individually, cannot be evaluated in a study using group administration of the test.

In a carefully designed experiment investigating the influence of color on the reactions of college women to ink blots, Siipola[4] discovered that the presence of hue-form incongruity created a conflict situation. Those blots having a color inconsistent with the dominant, form-suggested concepts showed striking differences from their achromatic counterparts with respect to reaction time and content. This would suggest that color does have a decided influence on responses to the Rorschach Test in certain instances.

Because any psychological test which is novel and ambiguous provides a mild stress situation for the subject, the Rorschach Test could reasonably be expected to give some measure of the degree of control ordinarily exerted by the subject when under mild stress. In view of the above findings regarding the influence of color on reactions to ink blots, the way in which the individual reacts to the colored Rorschach cards should prove particularly important in any evaluation of impulsiveness or lack of emotional control in social situations.

The present study is chiefly concerned with the establishment of reliable, objective signs of impulsiveness in the Rorschach Test with particular reference to the understanding of the normal superior adult.

Method

Two independent groups of college students were carefully selected from men living in Stanford Village, a housing project of Stanford University, in such a way as to insure that each member of a particular group was intimately acquainted with all other members of the group. The two groups of subjects consisted of 24 men in Building A and 22 men in Building B, hereafter referred to as Groups A and B, respectively. The close living quarters of these two buildings were conducive to occasional friction and dissatisfaction among the students. In a sense some of the group members were in a constant "stress" situation where intragroup frustrations gave rise to aggression, avoidance of participation in the group, and similar reactions.

In a preliminary session prior to the administration of the individual Rorschach Test, each subject placed himself and all other members of his group in rank order on ten traits including im-

pulsiveness. In this way it was possible to determine average rank values for each subject which could be used as criteria for comparison with the performance of each individual on the Rorschach Test. Further details of the methodological procedure used in the collection and analysis of the data are available in an earlier paper of this series[1].

Some evidence of the appropriateness of the criterion is present in the case history data obtained from interviews with the counselor of Group A. Individual M, who was judged the most impulsive by the group, was described by the counselor as being a very immature person who completely lost control of his emotions on several occasions with insufficient reason for doing so. Specific instances were mentioned by the counselor to substantiate his general impression. A second example illustrating the sensitivity of social judgments of impulsive behavior concerned individual J. It was noticed that J's rank changed considerably toward the end of the period in which the ranking data were collected. A later interview with the counselor, who was not aware of the ranking results, revealed that about half way through the ranking phase of the study J was involved in an argument during a bridge game which resulted in his impulsively upsetting the bridge table and walking out of the room. This suggests that judgments of impulsiveness made by an individual's associates are directly contingent upon the social behavior of the individual being judged.

Results

Three hypotheses concerning impulsiveness and the Rorschach Test were examined. The first two of these dealt with the predictive efficiency of isolated test scores. One of the more generally accepted hypotheses concerning indicators of impulsiveness in the Rorschach Test is the relationship between the number of color responses having definite form and those with indefinite form. A test of this hypothesis was made by placing all individuals in rank order according to the ratio CF:FC and then computing the correlation between this rank order and the criterion. This was done later with Group A as well as with Group B since the criterion data for A are more satisfactory than that for B. Three individuals in Group B gave no color responses and were not included in the test of this hypothesis. The validity coefficients were .42 and .07 in Buildings A and B, respectively.

A somewhat more inclusive hypothesis sometimes proposed is that indefinite shading responses as well as indefinite and formless color responses are indicative of a lack of emotional control. The specific hypothesis tested was that the degree of impulsiveness manifested by an individual in social situations would be directly related to the ratio (CF+cF +2C:(FC+Fc). The correlations between the Rorschach scores and impulsiveness were negligible for both groups, .18 in Group A and .03 in Group B.

The above results merely confirm the feelings of many Rorschach workers that, in general, consideration of a single aspect of the Rorschach only leads to misinterpretation.

In order to overcome the difficulties encountered by isolating single test scores, an extensive analysis was made of the test data for Group B without any knowledge of the data for Group A. It was proposed that *impulsiveness would be revealed in the Rorschach in one or more of a number of ways depending upon the particular personality structure of the individual.* One subject might disclose his impulsiveness by bursting forth with a spontaneous remark pertaining to color. Another might demonstrate a lack of control by using a predominance of formless concepts or color responses having indefinite form. In order to select variables which seemed to indicate an overreactivity to the ink blots, a large number of response characteristics was studied. The Rorschach protocols in Group B were compared with the criterion data prior to any analysis of the data for Group A in order to discover patterns of scores which discriminated the most impulsive from the least impulsive individuals. On the basis of this analysis eight signs were selected which seemed logically related to impulsiveness. These were:

1. immediate remarks provoked by the color
2. the relationship of CF to FC
3. the number of C in relationship to the other color responses
4. the relationship of (CF+cF+2C) to (Fc+FC)
5. the ratio of mean achromatic reaction time to mean chromatic reaction time
6. the mean time per response
7. the percentage of pure form responses
8. the number of FC in relationship to the other color responses

Each of these signs was given a weight according to its efficiency in predicting impulsiveness. It was necessary to lean heavily on a rational as well as empirical plan in weighting the signs to allow for any variations which might occur in a new group of Rorschach protocols. The weights given each of these response characteristics are presented in Table 1.

An "impulsiveness" score was computed for each individual by algebraically summing the weights of the signs checked on his Rorschach, each plus or minus signifying one point. All the subjects were then placed in rank order ac-

TABLE I
RORSCHBACH "SIGNS" OF IMPULSIVENESS

Weight	Sign
++	each card to which an impulsive, immediate (within 5 seconds) remark is made provoked by color
+	$2 \leq CF:FC < 3$
++	$3 \leq CF:FC < 4$, and $FC > 1$
++	2 CF and no FC
+++	$CF:FC \geq 4$, and $FC > 1$
+++	3 or more CF and no FC
+	1 C* and at least 1 FC
++	2 C and at least 1 FC
++	1 C and 1 or more CF but no FC
++	2 or more C and $FC \geq CF \geq C$
+++	3 or more C and $FC < CF < C$ (For each C beyond 3 add one additional plus to the base of 3.)
++++	1 C and no other color responses
++++	2 or more C and no other color responses (For each C beyond 2 add one additional plus to the base of 5.)
+	$(CF+cF+2C):(Fc+FC) > 1$
+	(mean achromatic reaction time): (mean chromatic reaction time) ≥ 1.5 and < 2
++	$Acht:Cht \geq 2$
—	$Acht:Cht \leq \frac{1}{2}$
+	$T/R < 20$
—	$T/R > 100$
—	$65 < F\% \geq 75$
— —	$F\% > 75$
—	2 or 3 FC but no CF or C
— —	4 or more FC but no CF or C

* Count Cdesc, Csym, additionals, and doubtful pure C as one half.

cording as their "impulsiveness" scores, those with the highest scores being at the top. The fit of these weighted signs to the data from which they were derived is illustrated by a correlation of .60 between the test and criterion data. This correlation, of course, offers no proof of the validity of the signs.

When tested on Group A the correlation between "impulsiveness" scores from the Rorschach and the criterion was .49, a validity coefficient significant at the .02 level of confidence.

Discussion

One thing seems to be quite clear. Since three of the eight signs of impulsiveness are of such a nature that it would not be possible to use them in the analysis of group administered Rorschachs, much vital information concerning emotional control is lost by use of the group Rorschach. In the present study the most important single indicator of impulsiveness was the presence of one or more spontaneous immediate remarks provoked by color. Consequently, Lazarus' conclusion that color does not influence performance on the Rorschach Test appears unwarranted.

In spite of certain limitations in the sign approach there is little doubt but that impulsiveness can be evaluated by the Rorschach Test if a number of variables are considered together rather than just one at a time. Further research directed at the more accurate weighting of the above signs is needed before they can be used with any confidence for individual evaluation. Future investigations concerning impulsiveness should also consider Beck's F+ per cent or some similar purported measure of intellectual control. The results of Williams' experiment[5] comparing the Rorschach Test with reactions under stress clearly illustrate the importance of form accuracy as a measure of control. A drop in form level in response to the colored ink blots might prove significant although a preliminary analysis of the present data gave negative results in this respect.

Summary

The use of single, isolated ratios in the Rorschach Test did not prove satisfactory for the prediction of impulsiveness as judged by the intimate associates of college men. This confirms the beliefs of many psychologists that a valid interpretative inference cannot be drawn from a score when it is abstracted from the whole configuration of responses and considered alone.

When a comprehensive analysis was made by considering a large number of plausible signs based on the interrelationships of variables as well as on their nature taken separately, the results were more promising. The verification of a hypothesis predicated on the assumption that impulsiveness is revealed in a number of different ways in the Rorschach Test, depending upon the individual personality structure, indicates that the trait of impulsiveness, or lack of control in social situations, can be evaluated by the Rorschach Test.

Bibliography

1. Holtzman, W. H. Validation studies of the Rorschach Test: Shyness and gregariousness in the normal superior adult. *J. clin. Psychol.*, 6, 1950, 343-347.
2. Klopfer, B. and Kelley, D. M. *The Rorschach technique.* New York: World Book Company, 1946.
3. Lazarus, R. S. The influence of color on the protocol of the Rorschach Test. *J. abn. and soc. Psychol.*, 1949, 44, 506-516.
4. Shipola, E. M. The influence of color on reactions to ink blots. *J. Personality*, 1950, 18, 358-382.
5. Williams, M. An experimental study of intellectual control under stress and associated Rorschach factors. *J. consult. Psychol.*, 1947, 11, 21-29.

RORSCHACH CONTENT IN HYPNOSIS: CHRONOLOGICAL AGE LEVEL REGRESSION[1]

MARGARET MERCER, PH.D. AND R. W. GIBSON, M.D.

Veterans Administration, Coatesville, Pa.

Introduction

The subject of this experiment is a 26 year old male college graduate who has partially completed the requirements for a master's degree in economics.

The patient is a divorced white male Roman Catholic who entered the hospital with the chief complaint of alcoholism. He is tall, overweight, and somewhat effeminate in appearance. He dated the onset of his symptoms to four years before admission while he was in the Army Medical Corps. His failure to advance and a poor marital adjustment led to excessive drinking. A vicious cycle of drinking, shame, and more drinking finally culminated in a dramatic but feeble suicide attempt which led to his discharge. He continued drinking intermittently. He had frequent violent quarrels with his parents and wife during the following years, but nevertheless was able to graduate from college and do some work for a master's degree. His marriage ended in divorce one month prior to admission.

His parents were both foreign born. The father, a steelworker, always rejected the patient and treated him with harsh brutality. He frequently called him a girl and said he was not his own son. The mother was overpossessive and showed her affection by overfeeding and pampering the boy. The maternal grandmother, a strongly religious and superstitious woman, showered him with affection and was often his sole protector against the cruelty of his father. An older sister, apparently following the father's example, treated him with disdain and frequently cruelly teased him.

At the age of five the patient witnessed parental intercourse. He was discovered by his father, beaten savagely, and threatened with castration. He adjusted poorly in school, being physically inferior to his classmates. He had skipped second grade and had difficulty in getting along with his studies. At ten he experienced an extremely traumatic event in the death of his grandmother. Since it closely followed one of his misdeeds which displeased her, he felt the full responsibility for her death. Shortly after he was seduced by an older girl. Though he fled in fright before experiencing orgasm, he began masturbating at this time. His first sexual intercourse at 17 was associated with strong guilt feelings and fear that he might contract a venereal disease. He dated many girls in a compulsive fashion, attempting to demonstrate his masculinity. Despite strong family opposition he married when 22 chiefly because he felt he "must make an honest woman" of the girl with whom he had had frequent sexual intercourse. The marriage was a happy one for the short period until his present illness began. In adult life he was approached many times by homosexuals and feared there might be something in his appearance which attracted them; he denies any perverse activity.

The dynamics of this man's personality can be only briefly presented, omitting much of the material on which the conclusions are based. He has a strong conflict between passive dependent needs and aggressive trends, apparently a reaction formation against strong castration fears. A rigid unreasoning superego makes the conflict especially difficult. Alcoholism provides gratification for oral dependent needs, a good medium for "he-man" behavior resulting in a form of punishment which appeases the demands of his superego.

In the course of psychotherapeutic interviews the patient was found to be an excellent subject for hypnosis. It was decided to give psychological tests to the patient under hypnosis by the method of Bergmann, Graham, and Leavitt[(1)].

The patient was given the Stanford-Binet Vocabulary, Goodenough Drawing, and Rorschach on three successive days when regressed to six, ten, and fourteen years of age respectively. Later

[1] Published with the permission of the Chief Medical Director, Department of Medicine and Surgery, Veterans Administration, who assumes no responsibility for the opinions expressed or the conclusions drawn by the author.

the Rorschach was given in the usual fashion.

When the patient was under hypnosis the Rorschach administration differed from standard procedures in that the examiner inquired "Is there anything else you see, anything else it makes you think of?" when the patient had apparently finished responding to the card. For this reason only the initial reaction time is reported for each card. The inquiry was conducted as soon as the patient completed his responses to each card. If the patient blocked, the therapist upon a signal from the examiner, said "It is all right to tell us what it makes you think of, you can tell us". The content of recurring or original responses was later investigated. At the close of each testing session the patient was told that he would not remember it. Each time the Rorschach was administered under hyphosis, the patient said he had not seen the cards previously.

FIRST SESSION: (Regression to six years of age)

The patient was hypnotized, told by the therapist that he could open his eyes and that he was now in first grade. The examiner was introduced as some one who was coming to the first grade to give the children some tests. The examiner then entered into the situation, talked with the patient about first grade activities, and started testing by saying she was going to ask him what some words meant. At the close of the testing session the patient was told that he would not remember it. Responses to the *Stanford-Binet Vocabulary* are as follows:

1. orange	1. you eat it
2. envelope	2. for letters
3. straw	3. for horses
4. puddle	4. water
5. tap	5. noise
6. gown	6. sisters wear it
7. eyelash	7. on your eye (shown correctly)

Six successive failures.

The drawing of a man produced at this session scores at five years three months according to Goodenough's scoring. It was the examiner's opinion that responses to the drawing and the vocabulary clearly established the patient's thinking at the six year level. The Rorschach was administered according to the procedures previously described.

RORSCHACH RECORD

Card I. Reaction time 20″

1. Birds flying.
2. Cats climbing up a tree.

Inquiry: 1. (Each side D is a bird. The head and wings are shown) 2. (Each side D is a cat. The center D is the tree)

Card II. Reaction time 15″

1. Dogs.
2. Flowers.

Inquiry: 1. (Usual dogs heads) Shape. 2. (Top red D) Because they are red.

Card III. Reaction time 7″

1. Men playing a drum.
2. Bowtie.

Inquiry: 1. Usual. 2. Shape.

Card IV. Reaction time 25″

1. A girl. She has pigtails. She is looking the other way.

Inquiry: W. (The lower center D is the neck, the lower side D's are the pigtails, and the upper side d's are explained as hair)

Card V. Reaction time 5″

1. A butterfly flying.

Inquiry: 1. W. Shape.

Card VI. Reaction time 22″

1. A statue of an angel.
2. A face, an old man's face. He has a beard.

Inquiry: 1. Top D. Shape. 2. Side D, usual. Shape.

Card VII. Reaction time 15″

1. A woman talking to another woman.

Inquiry: 1. Usual.

Card VIII. Reaction time 10″

1. A squirrel climbing on a tree.
2. A piece of meat.

Inquiry: 1. W. (usual animals. All center is the tree) 2. (Bottom pink) Because it is red.

Card IX. Reaction time 15″

1. A skull.

Inquiry: 1. (Center portion between lateral green) Because it has eyes. (When asked if he has ever seen a skull, says "In my prayer book")

Card X. Reaction time 10″

1. Two men shaking hands.
2. Flowers.
3. Spiders.
4. Two rabbits.

Inquiry: 1. (Large pink and center blue) 2. (All yellows) Like flowers at Easter time in church, yellow. 3. (Side blue) They have legs. 4. (Top grey) They are looking at each other.

Clearly some of the patient's responses are reflecting his religious training. He sees the statue of an angel in Card VI, Easter flowers in Card X, and the skull in Card IX has been suggested by the skull pictured in his prayerbook. The response to Card IV is probably of some personal significance. Questioning by the therapist later showed that this patient's sister, whom he strongly disliked, wore her hair in braids. It is interesting that this association is made to Card IV which frequently brings forth ideas of threat and insecurity.

SECOND SESSION: (Regression to ten years of age)

The same procedure was followed as has been previously described. This time the patient was told that he was ten years old. In answer to the therapist's questions he said that he was in seventh grade. This was correct. The patient had skipped second grade and his eleventh birthday occurred during the school year.

Responses to the Stanford-Binet Vocabulary.

1. orange	1. fruit
2. envelope	2. put letters in it
3. straw	3. dried hay
4. puddle	4. mud
5. tap	5. make a noise
6. gown	6. a dress
7. eyelash	7. you have it on your eye (showed correctly)
8. roar	8. make a loud noise
9. scorch	9. burn
10. muzzle	10. put it on a dog so he won't bite
11. haste	11. fast
12. lecture	12. don't know
13. Mars	13. a star
14. skill	14. to do things good
15. juggler	15. in a circus, he throws balls in the air

Six successive failures.

The drawing of a man produced at this session scores at ten years six months according to Goodenough's scoring.

RORSCHACH RECORD

Card I. Reaction time 20″

1. A coffin.
2. Curtains.

Inquiry: 1. (Center D) It is black. Shaped like a coffin. 2. (Side D) They are hanging around the coffin. They are dark.

Card II. Reaction time 22″

1. A steeple.
2. A dog.
3. Pieces of meat.

Inquiry: 1. (Upper cender d) Shape. 2. (Usual dog's heads) A collie. They are bushy. 3. (Top red) Because they are red.

Card III. Reaction time 5″

1. Two men. They are picking up a bucket.
2. Blood.

Inquiry: 1. (usual) 2. (Side red) It is red and it is running.

Card IV. Reaction time 25″

1. Half of a man.
2. This might be a cow, the head.

Inquiry: 1. Cut-off W. eliminating lower center D. Side D's are the legs and the small upper side extensions are "just something hanging to the man, his suspenders". 2. (Lower center D) It has eyes and horns. (2′ 10″)

Card V. Reaction time 4″

1. A bat.
2. Legs.

Inquiry: 1. W. Wings, head, feet, long ears. 2. (Side extension) A leg with no toes. A woman's leg because it is round.

Card VI.

Patient blocked. He was visibly upset, struggled to say something but could not bring himself to do so. The therapist assured the patient that it was not wrong for him to say what he was thinking of. The patient finally indicated the top D and said in a very low voice "when you go to the bathroom".

Card VII. Reaction time 15″

1. Two ladies talking.

Inquiry: 1. usual.

Card VIII. Reaction time 10″

1. Inside of a chicken.
2. A mouse.
3. A Christmas tree.

Inquiry: 1. W. Everything is different colors. 2. (Side D) It looks like a mouse. It is climbing. 3. (Top grey) Shape.

Card IX. Reaction time 17″

1. A skull
2. Two crabs.

Inquiry: 1. (As seen in Record I) It is white. There are two holes where the eyes are. 2. (Side orange) They have claws.

Card X. Reaction time 10″

1. A spider catching a bug.
2. Two people.
3. Rabbit's head.
4. Flowers, tulips.

Inquiry: 1. (Upper blue with green) Because it has lots of legs. 2. (Pink) They are dancing. 3. (usual) Nose and two ears. 4. As seen in Record I.

We have seen the association that caused the patient to block on Card VI. Responses to Card I, the first responses to Cards IV, VIII, and IX were investigated. It was learned that the patient's grandmother, who had lived with his family, had died before the patient entered 7th grade. She had been rather protective and indulgent towards the patient, he felt very strong affection for her, and was upset by her death. He says there are pictures of skulls in lots of places, and in holy pictures. These pictures make him think about when you die, you become all bones. You go to heaven unless you are bad. Then you go where the devil is. He feels afraid. He has seen one dead person, his grandmother. He later told the therapist that the experience had been terrifying and was followed by guilt, phobias, and night terrors.

Card IV has been seen as half of a man, his suspenders hanging down. This association relates to an episode in which the patient inadvertently came upon his parents when they were having sexual intercourse, was chased by his father and accused of snooping.

THIRD SESSION: (Regression to 14 years of age)

At this session the procedures described previously were followed. The patient was told that he was 14 years of age. When asked what grade he was attending in school he said he was in second year of high school.

STANFORD-BINET VOCABULARY:

1. orange	1. a fruit
2. envelope	2. a piece of stationery, a container for letters
3. straw	3. dried hay
4. puddle	4. sort of a hole in the ground with muddy water
5. tap	5. a noise
6. a gown	6. a dress
7. eyelash	7. hair on your eyelid
8. roar	8. a loud noise such as an animal makes
9. scorch	9. to burn
10. muzzle	10. what they put on a dog's mouth to keep him from biting
11. haste	11. go quickly
12. lecture	12. a teacher makes a speech
13. Mars	13. a planet
14. skill	14. to do something well
15. juggler	15. a performer (explains)
16. brunette	16. a color of hair, brown
17. peculiarity	17. something strange
18. priceless	18. very great value
19. regard	19. to know about a person or thing
20. disproportionate	20. not equal
21. shrewd	21. clever
22. tolerate	22. to get along, to understand
23. stave	23. don't know
24. lotus	24. a flower
25. bewail	25. crying
26. repose	26. lie down
27. mosiac	27. don't know
28. flaunt	28. show off

Six successive failures.

The drawing of a man produced at this session scores at 12 years six months according to Goodenough's scoring.

RORSCHACH RECORD

Card I. Reaction time 20″

1. Two eagles climbing.
2. A frog cut open.

Inquiry. 1. W. Each side D is seen as an eagle. The head, wings, and claws are identified. The claws are the two lower connections between the side and center D. 2. W. "Like in biology. Cut down the middle".

Card II. Reaction time 15″

1. Two bears heads. They must have been cut off. They are bleeding.

Inquiry: Usual bears' heads. Lower D is blood.

Card III. Reaction time 11″

1. Two men carrying something.
2. Pork chops.

Inquiry: 1. (usual) 2. (Side D) Shape and because they are red.

Card IV. Reaction time 28″

1. Some sort of alligator, just its head.
2. Boots.

Inquiry: 1. (Lower center D) Shape. 2. (Side D) Shape.

Card V. Refuses. The therapist told the patient it was all right to say what it made him think of. The patient then said "It is like some of those dirty pictures. Filth. They are not nice. That part has a bad name". It was impossible to determine the specific area indicated here.

Card VI. Refuses. The therapist again assured the patient that it was all right to say what it made him think of. The patient explained this as he had explained Card V. The examiner asked, pointing to the top D. What is that? The patient with some difficulty finally said "A statue". He would not explain anything further about the statue.

Card VII. Reaction time 8″

1. Two women talking.

Inquiry: 1. usual. The examiner pointed to the clasp detail asking "What is that?" The patient finally said "that looks like the same thing", then explained that it, too, reminded him of "these dirty pictures".

Card VIII. Reaction time 12″

1. Like insides of a cat.
2. Like squirrels.
3. A tree.

Inquiry: 1. W. Because of the different colors. 2. (usual) Shape. 3. (Top D) Shape.

Card IX. Reaction time 5″

1. Like insides of an animal, too.
2. A skull.

Inquiry: 1. W. These different colors. 2. As seen in previous records.

Card X. Reaction time 3″

1. A spider catching insects.
2. Witches dancing.
3. Like a girl with pigtails.

Inquiry: 1. Seen as on Record II. 2. (Pink D) 3. (Lower green) The usual rabbit's head without the ears is seen as the girl's heads while the usual worms are seen as pigtails.

This administration of the Rorschach is characterized by much more adequate handling of Card IV than has been seen previously, suggesting that repressive mechanisms have been established to deal with the threatening associations it has previously evoked. We have seen the associations which resulted in blocking on Cards V and VI. Investigation of the responses to Cards I, VIII, and IX dealing with the insides of animals are explained by the patient as related to his high school biology class. The return of the response concerning the girl with pigtails, this time to a small colored area, testifies to the strength of the association.

FOURTH SESSION: (Rorschach record taken in usual fashion without hypnosis)

Card I. 5″

1. This maybe is a vague outline of two eagles flapping their wings, climbing a post.
2. A woman's silhouette here.
3. Two snakes heads with their mouths open. Some of these things are embarrassing. Do you mind?
   E: Don't hesitate to say anything it makes you think of.
4. This may be a woman's breasts here.
5. On each side two animals heads with mouths slightly open.

Inquiry:

1. W. (each side D is an eagle) Wings, feathered claws. (because of shape only) This whole thing here (center D) might be a tree instead of a post.
2. Center D. Shape.
3. (upper clawlike extensions) Rattlers heads with jaws open.
4. (upper inner details) As if lying down and looking up.
5. (lower side) Shape. (2′6″)

Card II. 6″

1. Two Scottie dogs. Heads probably severed at the shoulders and bleeding, of blood.
2. Two dancing seals like you might see in a circus.
3. This might be a phallic symbol, one protrusion up here.
4. This white opening, as if you are looking in a deep cavern, brightly lighted.
5. These dancing seals also look like fresh pork chops.
6. This part here might be the profile of a primeval brutal face.

Inquiry:

1. Very hairy looking. A large flow of blood.
2. (top red) Shape.
3. (upper center d) Shape.
4. As if looking in. Very deep and brightly lighted.
5. The shape and red as if freshly cut.
6. Mouth open. Shape of the lower lip, brutal. 2′.

Card III. 6″

1. Two men, butlers perhaps, look as if picking up a basket.
2. Bowtie.
3. Two kidneys joined together with some membranes.
4. These things here are like raw hunks of meat torn out violently.

Inquiry:

1. Usual.
2. Shape, color has nothing to do with it.
3. I used to work in an autopsy room. Kidneys are not really that color but bloody.
4. Shape and color. (1′12″)

Card IV. 12″

1. This may look like female genitals.
2. Head of some animal like an alligator, scaly.
3. A body severed by a train, tattered and torn, the lower part of a body.
4. This part like a dog reclining on a little pillow here at the bottom.
5. The whole thing as if looking at a girl facing the other way, the back of her head with a bow on top and pigtails.

Inquiry:

1. This fissure here. The position. Lying down. The opening here as if there are feet. (W. omitting lower center D)
2. (Center D) Shape, rough looking skin.
3. A man with heavy shoes, torn in tatters. These (side d) are suspenders hanging down. The lower half of his body.
4. Looking up at something attentively. Just the head and neck. Shape.

5. Part in the center. Wavy texture like hair. (1'42")

CARD V. 10"

1. A bat flying.
2. A female leg, the lower leg without the foot.
3. Here as if someone probably a man were lying laid out with arms folded.
4. Phallic symbol.

Inquiry:

1. W. Shape.
2. (Side d) Shape.
3. (usual side) Arms folded. (demonstrates)
4. (top dd) Shape. (1'15")

CARD VI. 5"

1. This is clearly the female genital.
2. The male genital.
3. This looks like the statue of a mythological god, winged god with flapping wings.
4. Profile of an old revolutionary soldier, peaked cap and beard.
5. A bear skin.

Inquiry:

1. (Center D) Shape.
2. (Top D) Shape.
3. (upper part) Might be a bird god. Shape.
4. (side D) Shape.
5. (Lower D) Shape, looks furry. (1'16")

CARD VII. 7"

1. Two old women having a gossip session. Hair sticking up.
2. This fissure here, female genitals.
3. Here are two men, look like people dead for a long time, lips are shriveled and drawn back, looks as if smiling or laughing.

Inquiry:

1. Usual.
2. (Clasp detail) Shape.
3. (middle D) Like talking skulls. (1'16")

CARD VIII. 18"

1. Entire thing is like the inside of a body cavity during an autopsy, kidneys, stomach, and lungs.
2. Two guinea pigs climbing.
3. A fir tree, a Christmas tree.
4. Two animals looking up hungrily at this, some sort of food.
5. This fissure, female genitals.
6. Chest bones, the ribs coming out from the spinal cord.

Inquiry:

1. Different organs have different colors. You cut the cavity open and see these things.
2. Usual.
3. (Top D) A tree because of the triangular shape and greenish color.
4. As if animals (bottom pink) coming from a pit so you can just see their heads. Might be looking at these organs and trying to grasp them.
5. (Lower D) The color, the outline of buttocks, the fissure.
6. (Usual) Shape. (1'26")

CARD IX. 6"

1. The whole thing, the body cavity.
2. This section here is a skull. Eye cavities peering out from behind this stuff here.
3. This is like an optical illusion. The skull seems to contain one part. If you look at it another way the head swells up and moves up and down.
4. Here is a fissure again.
5. Pretty vague outline of lobster claws clutching at each other.
6. If you look at it this way the green part is like a stout woman washing clothes, a very pudgy face.

Inquiry:

1. W. Colors as before.
2. (Center portion between lateral greens) The greyish white color and the eye cavities.
3. If you look at it another way this changes and becomes larger, as you look at it. (Expands 2 to entire inner portion)
4. (Between halves of pink) Shape.
5. (Orange projections) Shape. These seem to have pinching movement.
6. (Green)
   1. add. (Lateral pink) Head of a man with his tongue sticking out. The way hanged people look. (1'42")

CARD X. 10"

1. This could be two people in costume doing a dance holding hands.
2. Yellow tulips.
3. Blue spiders chasing green beetles.
4. Vague suggestion of a fissure again.
5. A rabbit's head with something coming out of its eyes.
6. Baked potatoes.

Inquiry:

1. (Pink & blue) People don't usually dress that way. Hoods and long dress.
2. (All yellow) These (inner) are better shape. These (outer) are more like a cotton ball. Shape and color.
3. Don't need to be colored.
4. (In top grey) Shape.
5. Clear rabbit's head. Green stuff coming down
6. (Outer orange) Shape and color. (1'23")

The Rorschach record consists of 49 responses. The patient follows the ordinary gestalt qualities of the cards. Form quality is good. Popular responses are produced readily and all responses are easily followed. There is capacity for fantasy thinking and emotional responsiveness. This man's perceptual processes are clear and he is able to respond to promptings from within and to promptings from without. It is clear, however, that this is not a healthy personality. The amount and type of animal movement responses indicate a greater upsurge of instinctual drive than can be accepted and integrated. Content suggests that these instinctual drives have strong oral components. Fantasy thinking is present, much of it of a type to aid the patient in adjustment, but two

of the responses suggest disturbance in relation to ideas of sex and death. While the patient utilizes form and color in five responses, two are anatomical and one deals with food. Of the four CF responses, two deal with meat, one "raw hunk of meat torn out violently" and the other two use color to mark off areas of body cavities. The patient, himself, feels some need to explain his gory responses by pointing out that he used to work in an autopsy room. It will be remembered that when the patient was regressed to the fourteen year level, he gave responses concerning the viscera of animals and related them to his experience in his high school biology class.

While this patient has personality resources which would make it possible for him to utilize psychotherapy, the Rorschach record suggests that the patient's alcoholism is based upon deep-rooted insecurity resulting from long standing difficulties in relation to death and masculine adjustment. It may be difficult to reach him.

Summary and Conclusions

The subject of this study is a 26 year old male college graduate who has partially completed requirements for a master's degree in economics. The patient entered the hospital with the chief complaint of alcoholism. In the course of psychotherapeutic interviews the patient was found to be an excellent subject for hypnosis. The patient was given the Stanford-Binet Vocabulary, Goodenough Drawing and Rorschach on three successive days when regressed to six, ten, and fourteen years of age. At the close of each session, the therapist told the patient that he would have no memory for it. Later the Rorschach was given in the usual fashion.

1. Test data reflected changes at the various regressed age levels which were consistent with productions characteristic of the age level and followed the clinical data so clearly that there seems to be no doubt that true regression was established.
2. Some students of hypnosis believe that regression represents a genuine reproduction of an earlier period of life. Others hold that it presents elements of later years in the basic structure of the hypnotically induced age. Regardless which point of view is accepted, there can be little question that the responses repeated thru records taken at the various induced ages are of major importance to the subject. These repetitive responses indicate a deep-rooted personality disturbance related to sex, death, family relationships, and early religious training.
3. It is probable that the patient's traumatic experience related to his grandmother's death resulted in emotional attitudes which made it difficult for him to adjust to high school biology and the army autopsy room.
4. The sexual inplications obviously caused the patient to block completely on Cards VI and to the clasp detail on Card VII. The refusal on Card V and the explanation obtained suggests that the "black shock" frequently manifested in response to the card could profitably be investigated.
5. This type of study is, of necessity, limited to patients who are good enough hypnotic subjects to remain in the hypnotic state with their eyes open, able to respond to test material presented visually. It does indicate, however, the desirability of studying the significance of Rorschach content under conditions of induced free association such as may be obtained in amytal or pentothal interviews.

Bibliography

1. Bergman, M. S., Graham, H., and Leavitt, H. C. Rorschach exploration of consecutive hypnotic chronological age level regression. *Psychosomatic Medicine*, 1947, 2, 20-29.

A STUDY OF WECHSLER'S CONCEPT OF "NORMAL DETERIORATION" IN OLDER MENTAL DEFECTIVES

GERARD J. BENSBERG AND WILLIAM SLOAN

Lincoln State School & Colony

The present study is an attempt to supply an empirical check of Wechsler's hypothesis of "normal deterioration" in a group of mental defectives. Wechsler[1] *p. 28* states that a "result that might be expected from a study of the mental growth curve is that a child's IQ would tend to fall off as he grew older, and that this falling off would be more marked in the case of the mentally retarded than in the case of the normal or superior child."

Our particular interest is centered around the assumption that after age 25 or 30 there is a marked decline in abilities which is reflected in a decrease in IQ, according to Wechsler, if a constant CA divisor is used. To overcome this artifact, Wechsler has substituted a method for calculating IQ's which is based on the distribution of scores at each age and the calculation of IQ's around a fixed point (i.e. - 1 P.E. as IQ 90.) Such a procedure, claims Wechsler, "enables us to calculate IQ's which maintain the same meaning throughout the life of the individual."

Wechsler's data are based on a cross-section sample of population at different age levels. The present study is a longitudinal one, utilizing scores obtained when the subjects were between the ages 15-24 and comparing these scores with those obtained in later years. Such a procedure should overcome some of the possibilities of obtaining deviations from randomicity inherent in cross-sectional designs.

Procedure

In the Psychology Department of the Lincoln State School and Colony there is an index file which contains records of all the tests administered to the patients since their admission. From this file were selected the subjects to be used in this study.

Wechsler[1] has found that individuals tend to reach their maximum functioning level at about the age of 15 and that no noticeable decrease occurs until after the age of 24. Hence, a test administered during this age range was assumed to represent the maximum intellectual functioning, before any deterioration had taken place. Since the presence of any active organic process might influence the amount of deterioration over a period of years only subjects who were diagnosed as Familial or Undifferentiated were used. This diagnosis is based on case history, psychological tests, and a thorough medical and neurological evaluation. Because of the relatively high basal of the Wechsler Intelligence Test only subjects with IQ's of 42 or over were used.

A subject was chosen, then, only if he met the following criteria:

1. He received a 1916 Binet between the ages of 15 and 24.
2. He obtained an IQ of 42 or better on the 1916 Binet.
3. He is now between the ages of 30 and 55.
4. He has been diagnosed as Familial or Undifferentiated.
5. No active psychosis or organic process was observed.

The subjects were selected randomly from the file of the Psychology Department using these criteria. The age range of 30 to 55 years was divided into periods of five years each. Thirteen subjects were chosen for each age group, five were males and eight were females. This division was made in order to render a representative sample of subjects for each five year period from 30 years of age to 50 years of age. This gave 65 subjects, 25 males and 40 females.

As mentioned before, the three tests used to measure present functioning are: The Wechsler Intelligence Test, Form I; the Stanford-Binet Intelligence Test, Form L; and the Arthur Performance Test, Form I.

Thirty of the selected subjects had received two of these tests within the past year. For convenience these results were used. Some question might be raised as to the validity of these results obtained

by different examiners. All of the examiners were employed as competent examiners by the State and since several different examiners administered the initial tests, it will be assumed that no systematic error existed.

The order of administering the three tests was counterbalanced as far as possible in order to eliminate any practice effects. The Block Design Test and the Feature Profile of the Wechsler and the Arthur Tests contain some of the same material. If these tests were administered within a short time interval the same scores were used on those block designs which were the same on both tests and on the Manikin Test. This could not be done on the Profile because of the different divison of the ear on the Wechsler face and on the Arthur face.

The measures obtained in this experiment are: (1) the IQ on the 1916 Binet administered between the ages of 15 and 24; (2) the IQ on the Stanford-Binet, Form L; (3) the IQ on the Wechsler, Form I; and the EQ (Efficiency Quotient) on the Wechsler. The EQ is an individual's mental ability score (on the Full Bellevue Scale) when compared with the score of the average individual 20 to 24 years of age.

Results

Table I shows the basic results obtained. The first row indicates the mean 1916 Binet IQ for the various age groups from 30 to 55 years. These IQ's were obtained when each subject in the various groups was between ages 15 and 24 years. These scores are used as the basis of comparison for the IQ's obtained at present age on the other three tests. It is apparent that the different age groups were not constant with respect to initial IQ. There is a gradual and constant decrease with age. It would seem that random selection of institutional subjects results in progressively lower IQ's as age increases.

Several trends with respect to the other tests appear. Excluding the WB IQ, we find that the general decrease noted for initial IQ is present on the other three measures. While these measures are slightly higher than initial IQ at age 30, this difference disappears with increasing age. The WB IQ starts out much higher than all other scores and remains constant as age increases. This constancy is in accord with Wechsler's findings for discrete groups sampled at various ages.

Quantitative confirmation for these data was obtained by use of an analysis of variance. In order to obtain a score that would have a uniform meaning, a difference score was used. This was obtained by subtracting the IQ's of the Wechsler Bellevue, Form I (WB I), the Arthur, Form I (AI), the Stanford-Binet, Form L (SBL), and the EQ of the WB I (W-EQ), from the original 1916 Binet IQ. For ease of computation a constant was added to these scores to render them all positive.

A three classification analysis of variance was computed on these data in order to determine any significant differences between the variables age, tests, and individuals within each group. In the initial analysis of variance it was found that Tests were significant and Age was not significant. After a survey of the means it was felt that Wechsler I was the cause of this variation and the data were regrouped to compare the Arthur I, the Stanford-Binet L, and the Wechsler EQ against the Wechsler IQ.

Table I
Mean IQ for Different Groups

	Present CA				
	30	35	40	45	50
1916 Binet IQ*	55.9	55.7	54.4	52.3	48.8
AI IQ	57.6	57.5	55.1	51.6	46.0
SB IQ	60.9	57.5	57.4	53.3	48.0
WI IQ	67.8	69.1	69.8	69.4	68.8
W-EQ	61.0	58.2	58.1	53.4	49.8

*This measure was obtained when subjects were between ages 15 and 24 years.

TABLE II
SUMMARY OF ANALYSIS OF VARIANCE

Source	df	Sum of Squares	Mean Square	F
Age	4	97.8615	24.6453	
Tests				
3 Tests	2	218.287	109.1435	1.52
3 Tests vs WB IQ	1	9422.4628	9422.4628	130.842**
Interactions				
3 Tests x Age	8	65.5590	8.1948	
W/3 Tests x Age	4	971.2872	242.8218	3.3718*
Individuals	48	6628.1538	138.0865	
Remainder	192	13826.6924	72.014	
Total	269	31230.3039		

* Significant at 5% level of confidence.
**Significant at 1% level of confidence.

Table II shows the result of this analysis of variance. The remainder, being the best estimate of σ^2, was used as the error term. The three measures were found not to be sufficiently different from each other to be considered samples from different populations. The effect of age on the three measures was not significant, showing that with increasing age there is no systematic change in IQ, contrary to expectation.

The difference between the WB I and 3 tests is significant at less than the 1% level of confidence and they may be considered to have come from different populations. The interaction between WI, 3T, and age is significant at between the 1 and 2% level of confidence showing that the increasing age has a systematic effect on the Wechsler I IQ. By a survey of the means one can see that with an increase in age the difference between the WB I and the other tests increases.

DISCUSSION

The data in this study indicate that as the groups increase in age there is a general decrease in IQ on the three measures of intelligence with the exception of the Wechsler IQ. However, we must note the initial IQ's obtained on these subjects when they were between the ages of 15 and 24 decreased gradually as our samples increased in age. We find here that our random sample apparently selects successively less bright individuals as age increases. The basis for this might be the fact that those individuals who are brighter fail to remain in the institution and are gradually weaned out as age increases. Our sample of subjects in the institution, therefore, will be individuals who have successively lower IQ's as age increases. We must note that the measures of intelligence obtained by the Stanford-Binet L. the Arthur, and the Wechsler EQ are in close agreement with the original IQ's obtained on these subjects when they were between the ages of 15 and 24.

It is interesting to note that the Wechsler IQ's for all these groups remain fairly constant, and this is in accordance with Wechsler's findings. However, Wechsler did not have initial measures on his successively older subjects and it is possible, therefore, that his selection of subjects at the lower intelligence ranges for increasing CA, may have conformed to the selection of subjects in this sample. This is borne out by the fact that Wechsler (p. 112) states that his older subjects in the lower IQ ranges were selected from individuals who were referred to the Bellevue Hospital or from individuals who were in institutions for mental defectives. The question might be raised, therefore, as to whether or not Wechsler had a representative sample of older individuals in lower IQ ranges in his standardization population. It might be conjectured, consequently, that he did not sample those individuals who were not in institutions but who were mentally defective and therefore, failed to include in his older group and lower intelligence group those individuals who

were not in institutions. It is legitimate to make the assumption, therefore, that Wechsler's "normal deterioration" does not hold for the subjects found in this study.

The IQ's as obtained on other measures appear to be relatively constant and the much higher IQ's found on the Wechsler appear to be an artifact which may have been due to the inadequate sampling in his standardization population. All of this, presumably, may have been caused by the failure to control the intelligence of his standardization group at the younger ages. The present study, having controlled this factor, appears to cast considerable doubt on the validity of Wechsler's sample. While the conclusions to be drawn from this study must be limited to the present population, these findings may be interpreted to throw some question on Wechsler's entire assumption of so-called "normal deterioration". It must be remembered that while these findings are pertinent for only this population, it may be entirely possible that his assumption may be correct for the other ranges of the standardization group not included in this study. This question cannot be answered on the basis of the present data. It may be conjectured that in our population some deterioration may take place with increasing age, but it is not as great as Wechsler's deterioration curves indicate. Perhaps, at the lower levels of intelligence special curves of deterioration are needed. It may be that those individuals having less intelligence mature later and deteriorate at a slower rate than those individuals of average intelligence.

Summary and Conclusions

Sixty-five subjects, ranging in age from 30 to 55 were administered the Wechsler-Bellevue, the Arthur Performance, and the Stanford-Binet Intelligence Test. These IQ's were compared with the IQ's obtained on the 1916 Binet when the subjects were between 15 and 24 years of age. The following results were obtained:

1. Although there is a trend toward diminution of difference on the Wechsler EQ, the Arthur I and the Stanford-Binet L with increasing age when compared with the 1916 Binet IQ obtained in "peak" years, this is not a statistically significant difference.
2. The Wechsler I yields IQ's significantly different from the other three tests.
3. With increasing age the Wechsler I IQ's become increasingly different from the other tests and from the initial 1916 Binet.
4. The IQ's obtained on older subjects using the Stanford-Binet L, the Arthur I, and the Wechsler EQ show marked conformity to the initial IQ obtained on the 1916 Binet at ages 15 to 24.

It is suggested that the "normal deterioration" which Wechsler found for older subjects may have been an artifact due to faulty sampling, at least for the lower intelligence groups.

Reference

1. Wechsler, D. The Measurement of Adult Intelligence, (3rd. Edition), Baltimore: Williams and Wilkins, 1944.

JUDGMENT OF ADJUSTMENT FROM DRAWINGS: THE APPLICABILITY OF RATING SCALE METHODS[1]

GEORGE W. ALBEE AND ROY M. HAMLIN

The University of Pittsburgh, Western Psychiatric Institute and Clinic

Veterans Administration, Pittsburgh, Pa., and the University of Pittsburgh

Introduction

In a previous experiment[(1)] the authors have demonstrated that clinical psychologists can make judgments of adjustment from drawings with significant reliability and validity. In the previous study, the drawings-of-a-man-and- woman of ten carefully selected cases representing as wide a range of adjustment as possible were presented in the form of paired comparisons. Instructions to the clinical psychologists serving as judges were to decide, for each pair, which drawing-of-a-man-and-woman reflected the better adjustment. The reliability of judgment for fifteen psychologists was .97, and the validity coefficient was .65. These coefficients were significant at the 1% and 5% levels of confidence respectively.[2]

Certain questions about the validity of the pooled judgments arose due to (a) the small number of cases judged and (b) the assumption of a unidimensional continuum of adjustment. The previous experiment obtained mean scale values for each of the drawings which placed them along a continuum. The present experiment constitutes a check on the applicability of such a criterion scale as an aid in the judgment of adjustment inferred from drawings. If the criterion scale represents a valid continuum, and if psychologists can match drawings with drawings on the criterion scale with satisfactory reliability, then a criterion scale would be a useful adjunct tool in the interpretation of patients' drawings. While there may be specific, molecular factors which are important in the interpretation of drawings, it is our observation that many, if not most, clinicians use a molar, global or "intuitive" approach to this projective method. The use of a criterion scale, while not denying the importance of individual factors, would lend some objectivity to global interpretations.

Method

The previous study, using fifteen psychologists as judges, obtained mean scale values for drawings-of-a-man-and-woman which were presented in the form of paired comparisons. These scale values placed the drawings on a continuum, with unequal intervals, ranging from the drawing judged to reflect the best adjustment to the drawing judged to reflect the poorest adjustment.

Using these ten drawings as a criterion scale, the drawings of three groups of subjects were matched with it, one drawing at a time, for comparable adjustment. The three groups included (1) a group of 21 out-patient veterans diagnosed as schizophrenic, (2) a group of 21 out-patient veterans diagnosed as anxiety neurotics, and (3) a group of 30 veterans who were being treated at a V. A. Dental Clinic. The three groups were equated for age, sex, educational achievement and veteran status.

A drawing-of a-man-and-woman was obtained from each of these subjects. This variation of the original Good-enough test is in line with widespread clinical practice. After the drawings had been obtained, all identifying material was removed, and each drawing was coded for later identification. The 72 drawings of the three groups were then arranged in a random order for presentation to the judges.

The criterion scale was arranged in order on a table before each judge. Each judge received the following instructions:

"These ten drawings represent, in the judgment of a number of psychologists, a continuum of general adjustment for the individuals making the drawings. I am going to show you, one at a time, a num-

[1]Published with permission of the Chief Medical Director, Department of Medicine and Surgery, Veterans Administration, who assumes no responsibility for the opinions expressed or the conclusions drawn by the author.

[2]Clinically unsophisticated students in an abnormal psychology class gave judgments which did not differ significantly from the judgments of the clinicians. The mean judgments of 28 students showed a correlation (rho) of .90 with the clinicians' judgments; and a correlation (rho) of .41 with the criterion.

TABLE 1. MEAN SCALE VALUES AND STANDARD DEVIATIONS FOR THE THREE GROUPS

Group	*Mean scale value*	*Sigma*	p^{12}	p^{13}
1. Normal	1.63	.42	...	...
2. Neurotic	1.06	.48	5.54	...
3. Schizophrenic	1.14	.50	...	3.68

ber of drawings made by adult male veterans of comparable educational level and ask you to judge at what point along the criterion scale the drawing should be placed in terms of your judgment of the individual's general adjustment. It is not necessary that the drawings look alike. You may decide that a drawing reflects poor adjustment even though it does not look very similar to any of the criterion drawings. In this situation you should decide at what point along the continuum the individual should be placed."

The drawings of the three groups were then presented in random order. As the judges indicated the point at which each drawing should be placed, the scale-value of that point on the criterion scale was recorded. A total of four judges was used. The approximate time required for the entire group of drawings was less than an hour for each judge.

The four judges were split into two pairs. The average scale value for each drawing for one pair of judges was correlated with the average scale value for each drawing for the other pair of judges to give a split-half reliability figure. This value was then substituted in the Spearman-Brown formula to give the reliability of the four judges. Next, the mean scale-value representing the average of the four judges was determined for each of the three groups. Critical ratios between the three groups were computed to determine whether significant differences appeared using this technique.

RESULTS

The reliability for the four judges participating in the experiment was .89. This value is significant at the one percent level of confidence and indicates that the judges showed a high consistency in their judgment of where on the criterion scale the drawings should be placed.

The validity of the criterion scale depends logically on its efficacy in differentiating the three groups of subjects. This is indicated by the critical ratios between the mean scale values of the groups. The critical ratio between the normal group and the neurotic group was 5.54. The critical ratio between the normal group and the schizophrenic group was 3.68. Both of these values are significant at the one percent level of confidence. There was no significant difference between the neurotic and schizophrenic groups. These values are presented in Table 1.

The distribution of scale-values for the normal group showed negative skewness while the distributions for the psychiatric groups showed positive skewness. When a scale-value of 1.60, which separates the lower five drawings from the upper five, is used as an arbitrary cut-off point below which the individual is considered to be definitely maladjusted, 75 percent of the subjects are correctly placed. One-third of the "normal" group fell below this point while 14 per cent of the neurotics and 28 per cent of the schizophrenics fell above it. These values are presented in Table 2. The Chi-square for this distribution is 15.74 which is significant at the one per cent level of confidence.

TABLE 2. DISTRIBUTION OF THE THREE GROUPS WITH RESPECT TO A CUT-OFF POINT.

	N	*Above cut-off*	*Below cut-off*
1. Normals	30	20	10
2. Neurotics	21	3	19
3. Schizophrenics	21	6	15

SUMMARY AND CONCLUSIONS

The present experiment indicates that the use of a criterion scale as an aid to judgment of adjustment inferred from drawings lends encouraging reliability to

judgment. Further investigations of the validity of this approach are in progress. Our findings suggest that this method is effective in differentiating a normal group from two groups of neuropsychiatric out-patients. One possible explanation for the lack of difference between the neurotic group and the schizophrenic group may lie in the actual lack of difference in level of adjustment for the groups. The "schizophrenic" group, as was suggested above, was composed of out-patients, many of them in fair remission and holding jobs. It is likely that selection of a more "maladjusted" group of schizophrenics would improve the discriminative value of the criterion scale.

One further suggestion from the present experiment is the advantage which would derive from having several drawings at each point of the criterion scale. Despite their empirical agreement, our judges expressed difficulty making decisions where a drawing was markedly different from those of the scale. Research contemplated includes the establishment of a criterion scale with several sets of drawings of approximately equal values placed at intervals along a continuum. From present evidence we may conclude that psychologists can agree consistently with each other in inferring adjustment from drawings, and that these inferences possess a reasonable degree of validity.

Bibliography

1. Albee, G., and Hamlin, R. The reliability and validity of judgments of adjustment inferred from drawings. *J. Clin. Psychol.*, 1949, 5, 389-392.

THE HANFMANN-KASANIN CONCEPT FORMATION TEST AS A MEASURE OF RIGIDITY IN RELATION TO COLLEGE APTITUDE AND ACHIEVEMENT

RALPH D. NORMAN, CHARLES A. BAKER AND DONALD G. DOEHRING

University of New Mexico

Introduction

The basic postulation of this paper is that perhaps one of the factors which accounts for the difference between *predicted* achievement in college, (as indicated by the ACE test), and *actual* achievement, such as represented by grades in college, may be that of rigidity in thinking. Goldstein[3] says that rigidity consists of adherence to a present performance in an inadequate way. Kounin[6] mentions miscellaneous drives which result in rigid behavior, such as fear of failure, uncertainty as to consequences of acts, compulsive needs, inflexible goals, failures, frustrations, insecurity, low success expectations, etc. Werner[9], discussing secondary rigidity, says that this is an impairment of abstract thinking which occurs when tasks are too difficult. Rapaport[8, 5], in discussing rigidity in the Hanfmann-Kasanin Concept Formation Test, states that desirable factors for success are flexibility and persistence. Thus, rigidity may indeed have an adverse effect not only upon performance on this test, but also may occur when individuals are presented with difficult tasks, such as those which are necessary for scholastic achievement in college.

The Concept Formation Test seems to be particularly adapted to higher level thinking. Hanfmann and Kasanin[5] found that the ratio of the average score of schizophrenics as related to the average score of normals was 5:2 for college-level subjects, but only 3:2 for subjects of low educational level. Cook[1] found five factors in the Concept Formation Test to be related, in low and middle ranges of correlation, to ACE scores and grade-point average. (The present study differs from Cook's not only in experimental design, but also in that we have attempted to discover whether the test can detect rigidity and thus predict lack of success in college where the ACE

fails in prediction.) Fosberg[2], using superior subjects, considered *any* logical classification of the blocks to be correct, although in the present study the Hanfmann-Kasanin method was used, Fosberg's approach being considered as impractical in view of time required of each subject.

Method

It was decided that Analysis of Variance would be the best technique for discovering any significant results. Three variables—ACE, grade-point average, and Hanfmann-Kasanin test results were used. These three variables were treated thus: ACE scores would be factor A, with high ACE being termed A, and low ACE being termed a. Grade-point average would be factor B, with high GPA being termed B, and low GPA being termed b. There would then be four resulting experimental groups: *AB*, *Ab*, *aB*, and *ab*. Each group would be given the H K and results would then be compared by Analysis of Variance. *Ab* would be the crucial group, since it was assumed that a person with high intelligence and low grades might be exhibiting behavioral rigidity.

Forty subjects were selected; ten in each of the experimental groups. These 40 came from an original group of 200 students in elementary psychology for whom complete ACE and GPA data were available. The method of assigning the members to each group was to compare standard scores on the ACE with standard scores for the GPA. The former were obtained from a distribution for all University of New Mexico students; the latter were derived from the original sample of 200, which, it was felt, was fairly representative. Table 1 presents the comparative ACE and GPA standard scores for each of the four groups and forty subjects of this study.

In the *AB* group, the average was 1.3 sigma units above the ACE mean and 1.5 sigma units above the GPA mean. In the *Ab* group, there was a mean difference of 2.2 sigma units between the ACE and GPA. The mean difference between ACE and GPA in the *aB* group was 1.9 sigma units. In the *ab* group the averages were −1.3 sigma from ACE mean, and −1.5 sigma from GPA mean. Statistically, it might be said that the *ab* group as a whole was as far below the mean ACE and GPA as the *AB* group was above them; also that the *Ab* group with a differential of 2.2 sigma units between ACE and GPA was approximately opposite to the *aB* group with a differential of −1.9 sigma units.

The subjects as a whole, with one exception, were either freshmen or sophomores. The mean ages of the four groups were roughly comparable:

Table 1. Differential ACE and GPA Scores of the Four Experimental Groups in Terms of Sigma Units

AB Group	Diff. from ACE Mean in σ-units	Diff. from GPA Mean in σ-units	Ab Group	Diff. between ACE and GPA in σ-units
Subject 1	+1.4	+1.5	Subject 1	+2.0
2	+1.3	+1.5	2	+2.8
3	+1.0	+1.2	3	+2.1
4	+1.1	+1.6	4	+1.7
5	+1.5	+1.4	5	+1.9
6	+1.1	+1.9	6	+3.4
7	+1.4	+1.5	7	+2.3
8	+2.1	+1.4	8	+1.8
9	+1.2	+1.6	9	+2.0
10	+1.1	+1.8	10	+1.7
Mean	+1.3	+1.5	Mean	+2.2
ab Group			aB Group	
Subject 1	−1.2	−1.8	Subject 1	−1.2
2	−0.9	−1.5	2	−1.3
3	−1.5	−1.8	3	−2.4
4	−0.9	−1.6	4	−2.8
5	−0.7	−1.4	5	−1.6
6	−1.4	−2.1	6	−1.9
7	−1.8	−1.1	7	−1.7
8	−1.3	−1.0	8	−2.8
9	−1.7	−1.6	9	−1.4
10	−1.1	−0.8	10	−2.3
Mean	−1.3	−1.5	Mean	−1.9

(*AB*, 24.4; *aB*, 21.4; *Ab*, 21.3; *ab*, 20.3). There were 8 men and 2 women in *AB* and *Ab*; 6 men and 4 women in *aB*; 7 men and 3 women in *ab*. While it would have been highly desirable to have held the factors of college class, age, and sex constant, it would have meant searching far beyond the original sample of 200 subjects, since only 10 out of 200 (5%) fitted into any of the four desired ACE-GPA combinations.

The scoring method of Rapaport[8] was used. He mentions that in order to facilitate the completion of the test without undue time, at least one block should be turned over every five minutes. This method was followed in this study, i.e., if the subject had not made a grouping in a period of five minutes following the turning of the last block, another block would be turned over, and this was scored against the subject.

Results

Tables 2 and 3 show the results on the Concept Formation Test for the four groupings of subjects.

The only factor which tends to indicate significance, the ACE, does not permit rejection of the null hypothesis. To be significant at the 5% level, the F ratio would have to be 4.11. GPA and interaction of GPA and ACE show less variability than would be expected by chance factors alone. Since the analysis of variance tended to indicate a relationship between ACE and the H-K, a correlation was run between these two variables; this r was .36, significant at the 5% level. An insignificant r of .18 was found between GPA and H-K. However, the r's are attenuated by the fact that our total group comprised selected subjects.

The authors rescored the results, taking into account only the total amount of time consumed, and disregarding blocks turned over. Again, an analysis of variance was performed, with results shown in Table 4.

The results are about the same as shown for Table 3, except that the relationship between ACE and H-K fell further away from approaching signi-

Table 2.
Scores of Subjects in the Four Experimental Groups on the Concept Formation Test

(Score = Time + No. of Blocks Turned x 5)

	Subjects										
Group	1	2	3	4	5	6	7	8	9	10	Totals
AB	78	35	26	23	105	102	18	43	45	12	487
Ab	38	86	34	48	40	41	62	35	69	13	466
aB	66	131	24	20	46	107	124	47	58	35	658
ab	15	124	55	111	67	23	67	70	66	23	621

Table 3.
Analysis of Variance of the Data in Table 2

	d.f.	*Sum Squares*	*Mean Squares*	*F*
Total	39	43,790		
A	1	2,657	2,657	2.33
B	1	84	84	——
AxB	1	6.4	6.4	——
Exp. Error	36	41,043	1,140	

Table 4.
Analysis of Variance of Data on Time Only, without Penalty for Blocks Turned Over

	d.f.	*Sum Squares*	*Mean Squares*	*F*
Total	39	27,980		
A	1	1,030	1,030	1.38
B	1	150	150	——
AxB	1	22	22	——
Exp. Error	36	26,778	744	

ficance. Still a third scoring method was used, taking into account the statement of the *principle of solution* of the H-K by the subjects. A rather arbitrary scheme was devised as follows: If the subject stated the principle correctly, nothing was added to his score; if prompted once, 10 points were added; twice, 20 points; three times, 30 points; and if the correct generalization was never given, 60 points. This resulted in the following amounts being added to each group's total scores: AB, 180; Ab, 150; aB, 210; and ab, 170. The resulting score changes obviously did not appreciably increase the variability of the data among the four groups, so no further statistical analysis was made.

Discussion

Using the original (Hanfmann-Kasanin) scoring method with the Concept Formation Test, an analysis of variance revealed no significant differences among the four groups of subjects. We were thus unable to conclude on the basis of this method that rigidity (if this is measured by the test) accounts for discrepancies between ACE and GPA.

These results brought about a more critical analysis of the scoring method to discover whether it measured rigidity validly. In the original scoring, the subjects were penalized for each block turned, but this scoring might also have penalized those who were able to change hypotheses readily. Fosberg[2] found that there were over 50 logical groupings of the blocks into 4 groups, and such being the case, any particular subject may have struck one or more of these groupings before the "true" one. Thus he might have been penalized by the Hanfmann-Kasanin method, since those subjects who had used quite a few hypotheses were severely scored, and such scoring was certainly not measuring rigidity. In view of this a modification of the original scoring technique was made.

This new scoring method used only the total time taken for the "true" solution as an indication of rigidity in thinking, the rationale being that rigid subjects would take much more time than non-rigid ones. However, the re-scoring method yielded results comparable to the original scoring method.

There are several possible interpretations of our results: (a) The H-K Test may not have been valid for the measurement of rigidity in our design, since rigidity may also have been a factor in relative success with the ACE. That there is overlap between the two tests is indicated by the low significant r found in our selected sample. (b) Rigidity may play no important role in affecting GPA, since other factors enter into college success (age, socio-economic status, outside work, motivation, etc.). (c) Our design would not reveal anything significant if rigidity tended to affect *both* the ACE and GPA. (d) The GPA is not an uncontaminated variable, being affected by class status, type of course, college or department in which enrolled, etc. (e) The Concept Formation Test, as Hanfmann and Kasanin[5] point out, was designed primarily as a qualitative tool, and any quantitative results would be tempered by inaccuracies resulting from individual differences of the subjects. (Nevertheless they used the quantitative method in studying schizophrenics.) Also results of Hanfmann and Kasanin, and early use of the test by Ach, indicate that to some extent the H-K is an intelligence test.

Summary

The aim of this study was to discover the effect of rigidity on success in college work, using the Hanfmann-Kasanin Concept Formation Test as a measure of rigid thinking, with Grade Point Average and ACE controlled. An experimental design was drawn up using forty subjects: 10 had high ACE and high GPA; 10 had high ACE and low GPA; 10 had low ACE and high GPA; and 10 had low ACE and low GPA. An Analysis of Variance was performed, with the general result that there were no significant differences found mong these groups on the basis of scores on the Concept Formation Test. Lack of significance occurred whether a time-plus-block-turned-over score was used or a time score alone or a score which took into account the principle of solution. An r of .36, significant at the 5% level, was found between the

ACE and the Concept Formation Test, but no other significant r's emerged. On the basis of the results, no evidence was found that rigidity, if such were measured, is related to either the GPA or to the interaction of ACE and GPA.

Bibliography

1. Cook, C. H. Relation of certain factors of the Concept Formation Test to scholastic aptitude and to scholastic achievement. *Clark Univ. Bull., Abst. Diss. and Theses*, 1947, *14*, 78-80.
2. Fosberg, I. A. A modification of the Vigotsky Test for the study of higher thought processes. *Amer. J. Psychol.*, 1948, *61*, 558-561.
3. Goldstein, K. Concerning rigidity. *Char. & Pers.*, 1943, *11*, 209-226.
4. Hanfmann, Eugenia, & Kasanin, J. A method for the study of concept formation. *J. Psychol.*, 1937, *3*, 521-540.
5. Kasanin, J., & Hanfmann, E. Experimental study of concept formation in schizophrenia; quantitative analysis of results. *Amer. J. Psychiat.*, 1938, *55*, 157-166.
6. Kounin, J. S. The meaning of rigidity: A reply to Heinz Werner. *Psychol. Rev.*, 1948, *55*, 157-166.
7. Rapaport, D., Gill, M., & Schafer, R. *Diagnostic psychological testing*, Vol. I. Chicago: Year Book Publishers, 1946.
8. Rapaport, D. *Manual of diagnostic psychological testing*, Vol. I, *diagnostic testing of intelligence and concept formation*. New York: Josiah Macy Foundation, 1944.
9. Werner, H. The concept of rigidity: a critical evaluation. *Psychol. Rev.*, 1946, *53*, 43-52.

GROUP DISCUSSION AS A METHOD OF PATIENT ORIENTATION IN A NEUROPSYCHIATRIC HOSPITAL

JAMES O. PALMER, PH.D.

Washington University School of Medicine
St. Louis, Missouri

At a Veterans Administration hospital, a twofold need for orienting the newly admitted patient to the hospital environment was apparent. First, orientation was deemed vital so that the demands of hospital routine would not add further trauma to an already anxious individual, particularly so that his future adjustment to hospital treatment and care would not be hampered by an antagonistic or fearful attitude. This psychologically disturbed person, who had been unable to meet problems of adjustment outside the hospital, was now required to meet new demands in the form of hospital routine and from the procedures of medical and psychological study. Too often the patient would be thrust into the hospital regime and its regulations without a preliminary explanation of their meaning or purpose.

Second, orientation was felt required as an adjunct to the study of the patient and as a means of alleviating the task of his custody. The untractable patient presented a disturbing problem to the professional staff who were required to make a comprehensive psychiatric and medical study of each new admission. The ward staff were faced with the continuous care and supervision of a large group whose behavior was often unpredictable. Individual attention to the immediate anxieties of each patient proved to be an overwhelming task for the professional staff, even though they often attempted to calm the most disturbed and to answer as many separate questions as possible. The ward staff was by and large untrained to handle the confused queries of patients regarding the hospital, even if they were aware of the problem at all, and when so pressed, resorted to authoritative punitive discipline. The ward staff also quite justifiably considered the task of orienting the patient to be primarily a professional obligation.

Rationale

A discussion group of all newly admitted patients as an orientation procedure was suggested by the reported success of this approach for similar problems and by the author's own experience with group therapy. This method has been most fruitful where a common and immediate problem faced

most of the group members—as was the case for these newly admitted patients.

Several advantages were posited, *a priori*, for the proposed method: 1. A large number of patients might be contacted and observed by a greater number of staff in less time. 2. Factors attributing to the patients' adjustment to the hospital might be studied directly. 3. The patient might gain an initial impression of the hospital as offering emotional security. 4. At least, his current anxiety might be alleviated by the extra attention and support of a group. 5. Most important, the principal purpose of the group was to offer information to the patient about the hospital in a manner he could accept.

Several questions arose in advance concerning the group method: 1. Would there be enough administrative advantage to warrant the extra time and energy which the group would require of the staff? 2. When should a patient start attending a group? At first, there was some hesitation about introducing a patient into a group immediately after he was admitted to the hospital. It was recognized that there would be no way of predicting whether the group discussion would be advantageous or detrimental to a particular new patient, or whether a patient could contribute to the group or might prove to be a disturbing element. 3. Would these newly admitted patients feel free to discuss even such impersonal problems as hospital adjustment? Experience in group therapy has shown that it often takes some time for a group of patients to develop a sense of security with one another such that they are willing to express their feelings on a subject.

Sample

With the above purported advantages and questions in mind, a discussion group was organized, consisting of all newly admitted male patients on a locked receiving ward, patients who were not physically ill, and who did not require maximum supervision. No preliminary diagnostic classification was attempted.

The present report covers the first three months of the groups' activities. During this period, one hundred different patients attended twenty one semi-weekly sessions, which were conducted at regular hours on the ward. The usual group at any one session consisted of about 13 patients, four new to the group and nine who had attended at least one previous meeting. Only a few continued with the group for more than two sessions, over more than a two week period. Nearly all were new admissions, having been in the hospital less than one week. The group was conducted by the author and the psychiatric social worker assigned to the ward. The resident psychiatrists, nurses and attendants alternated in participation. The size of the group made it feasible for more than two staff members to attend—the best ratio of staff to patients being about one to three.

Procedure

Since in each meeting there was a large proportion of new group members, it was necessary to repeat continuously the purpose of the group and to make explicit the group structure. The group leader would begin by introducing himself and his position on the ward, and in turn introduce the other staff members. The nature of the group was stated in terms of a discussion of "what different patients expect from the hospital and what may be expected while at the hospital . . . of getting to know the people around the hospital and of them getting to know you." Each patient was then asked to introduce himself and to state how long he had been hospitalized here. The patients who had attended previous meetings were asked to describe the meetings further to the new patients during this introduction. This introduction period served to set a permissive atmosphere for the ensuing discussion; very often the patients began bringing up their questions as they introduced themselves.

As much as possible, the leader encouraged discussion of each question by as many different patients as possible rather than giving direct answers. Very often the nature of the question—and more important, the feelings back of it—became more explicit as it was discussed. Only if necessary was the question re-

ferred to the staff members and then to one most concerned with the particular part of the hospital routine under discussion. For example, if a patient "griped" about not being permitted to shave himself, he would be asked to discuss this further, and other patients would be encouraged to comment on it; other patients would attempt to explain this regulation; more important, they too would reflect their dissatisfaction, so that the patient who offered the complaint would not feel alone in his expression of hostility. After several patients had commented on the question, the ward attendant would be asked to explain his view as to the necessity for this procedure, and to state the time, place, and rules for shaving. Further discussion by the patients would be encouraged as necessary. Usually several such questions would come up simultaneously, under such general rubrics as "ward problems," "social service," "medical and psychological examination," and "treatment."

One of the main problems for the leader was how to balance the discussion so as to obtain optimum participation by all members of the group. Very often one or two patients would dominate the hour, while other patients would hang back. On other occasions, the entire group would be withdrawn and noncommital. A constant problem was the actively psychotic patient, interrupting a fairly reality-oriented discussion with his own delusions or flight of ideas. Common sense techniques seemed most effective in handling these situations. A particularly voluble and aggressive patient might be interrupted, by stating that other patients might be helpful to him with their comments. A silent group might be activated by suggesting topics brought up by other groups; the remarks of the psychotic patient were often helpful when the feeling tone behind the irrelevant content was reality oriented.

The principal cathartic agent used by the leader to relieve anxiety concerning hospitalization was the existence of the group per se. It was not only the permissive atmosphere of the situation, in which the patient was in effect told "you may say what you please," but more important was the fact that this group of individuals with common problems acted and felt together about them. Thus the anxiety of the patient did not remain restricted to himself, but was dissipated by the fact that it could be seen as a feeling of the group. This "displacement"—or dispersement of individual anxiety onto the group proved very effective. It was as if the patient could say "all this can't be so bad since every one else feels so." Very often this fact was noted by the patients themselves. The staff, of course, interpreted continually this relationship between individual anxiety and group feelings.

Results and Conclusions

1. *Changes observed in the behavior of the ward.* Staff members who were not involved in the group discussions independently noted such manifestations of changes as: a) the whole ward was quieter during this period than at any previous time, b) there were fewer aggressive outbursts by the patients either between themselves or with the staff, c) patients were more observant of ward regulations, presented fewer complaints about these regulations, and cooperation in ward work improved, d) patients tended to carry on the group discussions autonomously as a method of solving inter-patient relationships. For example, one group of patients was overheard discussing how to gain the cooperation of a new patient in the ward work, and again how to support a particularly depressed patient whom *they* feared might commit suicide. From such observations it may be concluded that there was a general improvement in the patients' attitudes toward the hospital and particularly in their sense of responsibility for themselves.

2. *Attitudes toward treatment.* Many of the patients who had entered the hospital with extreme hostility and suspicion toward hospitalization would, after attending the group meetings, begin to inquire as to the possibilities of treatment, particularly psychotherapy. The demanding patient, who customarily pressed for immediate discharge or for complete ground privileges, would,

instead, express concern for solution of personal problems, and request interviews with the therapist. Other patients, the non-commital evasive individual or the recluse, frequently became more vocal, more willing to establish therapeutic relationships, and to face individual problems as well as immediate frustrations. Such results with this latter type of patient were particularly encouraging, since their response in the group discussion itself did not always seem significant.

3. *Changes in the work load of the staff.* Admittedly, organizing the group discussions and participating in the meetings was an additional task for the staff. In at least one instance, a doctor had to be relieved of other duties for this purpose. On the other hand, the attendants agreed that the energy required of them in getting patients to the group and attending the group themselves was much less than that previously required in disciplining unoriented patients. In general, most staff members found the additional effort required in conducting the group to be rewarded by the improved attitudes and behavior of the patients.

4. *Regarding the diagnostic possibilities* of these discussion groups, the extra hours of observations occasionally yielded new clues as to his personality but in most instances merely corroborated the more comprehensive evidence from other investigations. The most helpful angle of the group observations was in prediction of the reaction of a patient to such activity programs as occupational therapy or recreation groups. Frequently, a sounder estimate of the patient's reactions to psychotherapy was possible.

5. Perhaps the most striking result of this experiment was the marked *improvement in staff-patient relationships*, particularly attendant-patient relationships. Much of the hostility which many patients felt toward the discipline imposed by the attendants came out in the discussion and was handled directly in these meetings, on a realistic basis. In several instances, the particular attendants whom the patients most feared and resented were invited to the meetings on the request of the patients, and their differences were thrashed out at that time. By this method, the attendants—previously unaccustomed to dealing with the patients in terms of their feelings and emotional reactions—became less threatened by the personal attacks of these patients and were more aware of which of the patient's demands were realistic and which were projections of needs for attention, expressions of general hostility, etc. The professional staff learned to consider the patient as a member of a social environment, namely the hospital ward, became more aware of the daily pressures on the patient, and were reassured of the effectiveness and importance of their administrative "paper work." Many times the complaints of the patients against the attendants could be handled administratively by the physician. The professional staff also realized how extensively this daily routine affected the immediate reaction of the patient as he entered the hospital—thus affecting the physicians' diagnostic impression.

While both the ward staff and the professional staff became more aware of the personal problems of "being a patient," the patients, in turn, began to realize that "after all, the doctors are only human" and "maybe we should take into account the feelings of the attendants." These points were frequently raised in the discussion between the more overtly hostile patients and the more conventional individuals. For example, a patient complaining of neglect by the doctor, e.g., "I've been here ten days, and have only seen my doctor twice, and then for only fifteen minutes," might be urged by the staff to try to contact the doctor and inform him of his feelings. The patient might reply that the doctors are always out, could not be interrupted or that nurses and attendants discouraged his bothering the doctor. Other patients would then begin pointing out the doctor's side of the picture—his multiple load of duties, his responsibilities, etc. More often it was necessary for the staff to stress that the doctor needed as much communication from the patient as the patient was willing to give during these first weeks.

The effects which this group had on the cooperation between staff members

and on their relationships with patients was demonstrated in a project which both staff and patients undertook—namely, the composition of an "orientation manual." The need for such a manual was first voiced by the patients, and the information it contained was based on the questions which had been brought up in the group. As the staff gathered and edited the contributions, the differences concerning "what a patient ought to know," and consequently the differences in attitudes toward patients among the staff, became clarified. Through the discussion among the staff about these differences, a more unified understanding of patients needs was achieved.

THE NEGRO PATIENT IN PSYCHOTHERAPY*

RALPH W. HEINE

V. A. Mental Hygiene Clinic, Chicago

It is not surprising that the particular problems arising in psychotherapy with Negro patients have received so little attention. As a racial minority which has been low on the socio-economic scale, most private sources of psychotherapeutic help have been virtually closed to all but a very few Negroes. Now, however, in a large number of public mental hygiene clinics Negro patients are being treated in increasing numbers. Hence an early attempt should be made at least to identify some of the interpersonal problems characteristically arising in the course of psychotherapy with Negro patients.

No attempt will be made here to generalize on the character structure of Negroes as a group for the writer, at least, has never found characterological studies of entire groups of specific help in working with one member of that group. Rather, an attempt will be made to describe one or two of the barriers to a satisfactory treatment relationship which must almost inevitably arise when the patient is a member of a minority group which has very frequently if not uniformly received prejudicial treatment from a majority group with which the therapist is identified.

The chief barrier having its focus in the patient can be regarded as emerging from difficulties of the Negro patient in differentiating between what might be called the primary self-picture which is a product of his direct, intimate interpersonal experience with significant members of his family, his friends and his close acquaintances within the minority, and a secondary self-picture which emerges from experience as a member of a racial minority having a number of somewhat formalized contacts with the discriminating majority.

Stereotypic patterns of behavior can be and are evoked by members of any minority group. Since we all are members of minorities of one kind or another, we have all encountered situations where our characteristic personal qualities are overlooked while other qualities associated with a stereotype are attributed to us. For most of us such minor or even gross misjudgments of our personal characteristics are only annoying and sometimes even amusing but not too often of lasting importance. For most Negroes, however, it is a matter of primary importance and an almost constant source of tension and hostility It is not unrealistic for Negroes to assume on initial contact that no white person will assess impartially their traits, attitudes and abilities. If Negroes make a practice of covering up in relationship with whites, it is a realistically oriented defense mechanism and a useful tool in maintaining emotional equilibrium in social situations which are frequently threatening. In the interests of main-

*Reviewed in the V. A. and published with the approval of the Chief Medical Director. The statements and conclusions published by the author are his own and do not necessarily reflect the opinion or policy of the Veterans Administration.

taining reasonably comfortable social relations with whites, Negroes must adapt themselves in one way or another to a code of behavior that has been created largely by the white majority.

Stereotypically and traditionally, Negroes have been regarded by some of the white majority as being among other things loyal but passive, polite and jolly, sometimes irresponsible and as having no thought of competing with whites for prestige and status. However contradictory the beliefs of the whites and the actual feelings of the members of the Negro minority, it was and is a self-protective measure for Negroes to accommodate to this pattern to some extent at least. As more and more educational and professional advantages become available to Negroes (de jure at least) and as artificially created labor shortages (as occurred during WW-II) made desirable jobs available to Negroes at all levels of skill and ability, competition with whites has become more possible and more open, yet patently is not on an equal basis.

Accordingly, it is reasonable to anticipate that a poorly educated Negro or one who was reared in an area with a high level of prejudice expressed in legal restrictions and a double-standard educational system would be more likely to relate to a white therapist in a manner compatible with the traditional expectations of the majority group. He would be passive, polite, and essentially non-communicative; he would see the therapist not as a collaborator but at best as a sponsor to aid him and defend him in his contacts with the white world. The notion of communicating real feelings or interpersonal problems "in front" would be difficult for him to accept. On the other hand, the well-educated Negro who competes more openly for prestige goals in the culture and has attained some measure of success eschews the traditional pattern of inter-action. He nevertheless has no assurance of being accepted by the white therapist and hence is afraid to communicate real feeling. Rather, he substitutes social problems for personal problems.

There is really only one barrier having specific reference to the treatment of Negro patients which has its focus in the therapist and that, of course, is the therapist's[1] unresolved prejudice in the racial area. At the same time prejudicial attitudes may be expressed in more than one way as will be described below.

While for the Negro conscious covering up and conscious accommodation to a degree to the expectation of the white majority may have some utility in interracial relationships generally, it clearly has no utility in a psychotherapeutic relationship. Since no psychotherapist can treat the "Negro problem" he cannot treat a Negro who represents his problems as lying exclusively in his racial membership, any more than he can treat any person who persistently assigns the causes of his discomfort to people other than himself. A psychotherapist can intervene helpfully only in the lives of individuals experiencing problems in living which are to a large measure a product of that individual's distorted perception of his own needs and of his capacity to satisfy those needs within the real limitations imposed by his milieu. This applies to Negroes or members of other minorities quite as fully as to any other group of patients.

It was suggested earlier that every Negro has two bases of self-referent feelings and ideas—the attitudes expressed by significant adults during his rearing and the attitudes expressed by the white majority. Clearly these two influences

[1]Where the word "therapist" appears it means "white therapist." Where Negro therapists treat Negro patients many of the problems described here may still arise although on a more selective basis. It is likely that a dark-skinned, lower-class Negro would feel no less defensive with an educated, upper-class Negro than with a white therapist. To some extent Negroes who have professional status and move comfortably in a status-equal relationship with white colleagues may be identified in the clinic atmosphere with the whites rather than as belonging to the Negro minority. Furthermore, while the Negro therapist may not be thought of as "white" by other educated Negroes, he may have quite as much difficulty in establishing a treatment relationship because of competition and because the educated Negro may feel assignment to a Negro therapist where there is a mixed staff is derogatory or prejudicial treatment. Altogether, while the Negro therapist may not have precisely the same problems with Negroes as does the white therapist, his problems in establishing a satisfactory treatment relationship are non-the-less difficult and challenging.

are not totally independent but rather closely integrated one with the other. If a Negro experiences salutary relations with significant adults as a child and therefore has high esteem for himself, his inter-racial contacts will have a relatively less destructive effect upon him. Where, however, early relationships have been conflictual with consequent low self-esteem, inter-racial relations are extremely threatening and become the source of further self-derogation. We may reasonably assume that it is from the latter group that candidates for psychotherapy are drawn.

Among this latter group, therefore, are the people who are least able to look at themselves with an inquiring attitude and for whom the manifest difficulties encountered in inter-racial relation are the readily available basis of a highly effective defense. Depending upon the particular role the patient assumes in managing inter-racial relations, the patient may use the therapy hours in making more or less subtle pleas for sponsorship of the white therapist or as a forum in which the "Negro problem" is endlessly discussed in all possible ramifications. Usually the therapist is tacitly or pointedly excluded from the fraction of the majority group holding prejudiced attitudes. Obviously neither of these uses of the therapeutic hour do much to advance the patient's understanding of himself or his personal problems. One might ask: How does this defense present a problem different from that of any patient who uses therapy hours to describe the iniquities of the world and of the people in it? There are at least two differences: First, the white patient commonly describes the iniquitous world in terms of himself alone whereas the Negro patient can and does quite reasonably generalize to all Negroes. Thus, where one can hope eventually to help the white patient to see that his reaction is not entirely shared by others, with the Negro patient this is not so easy. Second, (and related to the first) one assumes with the white patient who sees all his problems as emerging from the faults of others that the significant people he describes are probably not quite as he describes them. When this assumption is borne out it becomes a basis for re-directing the patient's attention to himself. With the Negro patient one is often reasonably sure that various interpersonal situations with whites are quite often exactly as he describes it. Thus the therapist may seriously disrupt his rapport if he interprets as wholly projection or displacement what is literally true.

In this connection, perhaps, the role of the therapist's attitudes should be brought more into focus. Let us assume that the therapist has successfully resolved the most obvious conflicts with respect to Negro patients and experiences no conscious discomfort in working therapeutically with Negroes. There still remains, however, the considerable danger of identifying with the patient and feeling too sympathetic to be of assistance or of retaining some remnants of the prejudices of the majority and feeling too guilty to be of help.[2]

Such feelings on the part of the therapist can only lead to his communicating anxiety to the patient and thereby strengthening the defense described above. In either case the therapist is likely to be led easily into an attempt to treat the "Negro problem" rather than the particular problems of the particular patient with him—a patient who, in addition to whatever realistic and neurotically determined limitations he must cope with, has the additional realistic limitation of possessing a colored skin in a society which entertains prejudicial attitudes toward such a skin color. It is this limitation imposed by skin color which the white therapist may find fully as hard to accept as the Negro patient. Yet if the therapist is to be of help to the Negro patient he must acknowledge the status quo rather than joining with the patient in phantasy concerning the ideal world or in commiseration with respect to that which exists.

A moment's reflection will lead one to realize that from the viewpoint of individual emotional adjustment, being a

[2]See ADAMS, W. The Negro Patient in Psychiatric Treatment, *American Journal of Orthopsychiatry*, Vol. XX, No. 2, Pages 305-310. In this paper, published after the present paper was completed, some of the same points are discussed but with a somewhat different emphasis.

Negro is a handicap primarily in the area of competition with members of the economic white majority for social and prestige goals. It is no handicap in all other areas of interpersonal experience except insofar as the hostility arising out of unsuccessful and frustrated competitive efforts are displaced to the more intimate primary intra-racial relationships. It is in relation to this situation that the white person is likely to find his role as therapist and his role as citizen seemingly irreconcilable.

As a citizen the therapist may be active in efforts to reduce the frustration which Negroes as a group experience in competing for the economic and prestige goals of this society. As a therapist he must not only be alert to instances where patients are, perhaps, seeking discrimination by competing for the goals for which they are not suited intellectually or emotionally, but he must also often take the view that his Negro patient is being unrealistic in competing for goals for which in terms of personal talents and capacities he is quite suited. Put another way, the therapist must help his Negro patient to see that active competition for many of the prestige goals of the society will be successful only through sacrifices and compromises not demanded of white people.

The handling of the therapeutic problem discussed above is of course dependent on the therapist's typical manner in the therapeutic relationship and on his techniques. That is, if as suggested here the preoccupation of the Negro patient with racial issues is to be regarded as a symptom rather than as a problem properly within the scope of psychotherapy, the therapist would deal with this as he would with any other symptom.

The assumption here is that the need to relate to the therapist only in terms of the difference in race membership or in terms of racial problems is a reflection of unresolved interpersonal problems rather than the primary interpersonal problem. Thus, one may interpret, reflect, reformulate, direct through questions, or use whatever method his training and experience calls for in helping a patient to attend to his own interpersonal problems rather than to an incapacitating or troublesome symptom. As with any other symptom the psychotherapist accepts the patient's account of the discomforts arising out of the symptom but as soon as possible ignores the symptom and leads the patient to explore his feelings and attitudes concerning himself and his relationships with the persons who have had the most significant role in shaping his life.

However, to dissolve this defense, the therapist must clearly communicate to his patient that he (the therapist) is interested in him only as an individual. This, we have said, is possible only if the therapist really does see Negro patients as individuals and not as symbols or as representatives of the racial minority. The therapist must himself have reached some stable resolution of his attitudes toward the Negro minority thereby freeing himself to deal with the particular problems of the particular Negro patient seeking therapeutic assistance. When these barriers to a salutary treatment relationship have been overcome, the treatment of a Negro patient should differ in no important respect from the treatment of any person in this culture.

AN ATTEMPTED EXPERIMENTAL VALIDATION OF THE THEMATIC APPERCEPTION TEST

JAMES S. CALVIN
University of Kentucky
and
LEO C. WARD
Veterans Administration, Atlanta, Georgia

Psychological tests of the projective variety seem to have demonstrated their usefulness for psychiatric diagnosis. It appears, however, that their usage has spread beyond this field, so that now many psychologists, and others, apparently believe that projective tests are useful in assessing the personalities of other, non-psychiatric or "normal" individuals. Is this belief justified? The supporting evidence actually seems scant. This paper is the report of an attempt to explore this problem—the problem of the usefulness of the projective techniques in the prediction of normal behavior.

More specifically, the problem may be stated in this way: Can the results of projective tests predict how individuals will behave in situations in which they are placed? To answer this question, of course, it is necessary to devise a situation and then measure in some way the behavior of the individuals in this situation. These situation tests may then be considered as criteria for validating the projective tests.

Procedure

In this exploratory study, only one situation test was used. This was accomplished by means of the following apparatus and procedure. Ten light bulbs were mounted on a vertical board. Various combinations of five lights each were presented to the subjects, one combination at a time. The subject was asked to predict, each time a light combination was presented, whether a buzzer would follow. He was also asked to indicate his judgment of certainty or confidence for his prediction on the following scale:

5—for "dead certain"
4—for "almost certain"
3—if he considered it "a good bet"
2—if he would describe his feeling as "a strong hunch"
1—if he were "uncertain, but had a feeling"

In a previous experiment, not concerned with the present problem, it had been observed that, in the light-buzzer situation just described, individuals differ widely in their statements of certainty. Some persons fairly consistently used the 5 and 4 positions on the scale, indicating rather high certainty for their predictions in this situation. Others used lower regions of the scale. In the present experiment, prediction was actually a chance matter. Each light combination was followed by the buzzer on fifty percent of its presentations and it was *not* followed by the buzzer on fifty percent of its presentations. The total number of presentations for each subject was 120. Order of presentation of the trials on which buzzer was presented was random.

Ninety subjects were used in this light-buzzer situation. Of these, two extreme groups of ten each were selected for further study. One of these groups gave high certainty ratings; the other gave low certainty ratings. Mean certainty ratings for each of these groups are shown in Table I. Since these are selected extreme groups, there is no overlapping with regard to degree of certainty.

Table I
Certainty Ratings Of The Two Extreme Groups

Group	Number of Subjects	Mean Certainty Rating on First Trial	Mean Certainty Rating for 120 Trials
High Certainty Group	10	4.5	4.1
Low Certainty Group	10	1.0	2.2

The rationale for this admittedly exploratory experiment is somewhat as follows: It seems plausible to expect that some of the differences between the high certainty group and the low certainty group are personality differences. If this is true, it should be possible to detect these differences with a technique which measures personality. Under these assumptions, the Thematic Apperception Test (TAT) was administered to the twenty subjects comprising these two groups. The subjects' stories for each of the twenty test cards were recorded by means of a wire recorder and a concealed microphone. The subjects were unaware that their stories were being recorded verbatim.

Results

The TAT stories were analyzed in several different ways.

Analysis of Needs: First, an analysis was made in accordance with the need system described by Murray[6]. Scoring for these variables is a matter of judgment and it is difficult to be objective about this scoring. This is a probable source of inaccuracy. An attempt was made to curtail this inaccuracy by scoring only for manifest needs or traits and not going beyond the characteristics manifested to look for latent need or traits, and secondly, by requiring agreement between the two investigators before recording any score. This scoring was done "blindly". In other words, the investigators did not know which subjects were in the high certainty group and which subjects were in the low certainty group until scoring had been completed.

The results of this analysis are shown in Table II. Mean frequency of occurrence of each of the personality variables listed in Table II was obtained, and percentage of subjects in each group exhibiting each variable was computed. These comparisons are shown in Table II. Most of the differences between the high certainty group and the low certainty group fall short of statistical significance. Three of them, however, are significant at the five percent level or better. These

Table II

Personality Variables From TAT Stories of High and Low Certainty Groups

TAT Variable	Mean Frequency of Occurrence		Percentage of Subjects	
	High Group	Low Group	High Group	Low Group
Abasement	0.60	0.30	30	30
Achievement	2.90	3.60	80	100
Acquisition	0.60	0.80	30	30
Affiliation	0.50	1.50	30	60
Aggression	3.30	2.10	90	80
*Autonomy	0.70	0.10	50	10
Blamavoidance	1.20	0.60	70	50
Cognizance	0.20	0.20	10	20
Counteraction	0.40	0.10	30	10
Deference	0.10	0.10	10	10
*Dominance	1.40	0.70	90	60
Exhibition	0.50	0.80	30	40
*Harmavoidance	1.80	0.90	100	70
Nurturance	3.50	3.20	100	100
Play	1.50	1.90	70	70
Rejection	1.70	1.30	40	70
Sentience	0.80	0.60	40	30
Sex	3.00	1.50	70	80
Succorance	1.50	1.50	80	90
Understanding	1.00	1.50	50	80
Anxiety	0.30	0	30	0
Disjunctivity	0.90	0.20	40	20
Emotionality	0.50	0.60	40	30
Endurance	0.50	0.20	30	30
Endocathection	0.10	0.20	10	20
Intraception	0	0.50	0	20
Intensity	0.10	0.40	20	20

*Difference between high certainty group and low certainty group significant at the .05 level.

are autonomy, dominance, and harmavoidance. Of course, it is possible that these are only chance results. In any analysis of a large number of variables, a few differences should turn out to be significant at the five percent level or better from the operation of chance factors alone.

Analysis of Language: Next, an analysis of language usage, to determine the number of terms indicating uncertainty, vagueness, and so on, used by each subject, was made. This sort of analysis has been described previously by Balken and Masserman[1]. Frequency of uncertain statements, such as "may", "might", "could" were counted for each group. Similarly a count of vague modifying statements—"sort of", "kind of", "or something", "or anything", "and everything"—was made. The results of this analysis are shown in Table III. Neither of these differences is statistically significant. The fact that both these differences are in the same direction may add some support to a conclusion that the low certainty group uses vague and uncertain language to a greater extent than the high certainty group. However, when uncertain language is combined with vague language, the difference between the high certainty group and the low certainty group approaches but still falls short of statistical significance.

It might be expected that a person who tells long stories would use more vague and uncertain terminology than a person who tells short stories, simply because he uses more words. For this reason, a ratio of the frequency of these kinds of verbal usage to the total number of words used by each subject was determined. These ratios were computed and the means for the high certainty group and the low certainty group were compared. Again, the results indicate greater relative usage of vague and uncertain language by the low certainty group, but do not attain statistical reliability.

Analysis of Outcomes: Finally, the outcomes of the stories were rated with regard to their desirability. Were the outcomes favorable, unfavorable, or questionable? The results of this analysis are shown in Table IV. The high certainty group had fewer questionable or neutral outcomes and a greater number of unfavorable outcomes. The last of these differences is statistically significant at the 5 percent level.

TABLE III
LANGUAGE USAGE BY THE HIGH AND LOW CERTAINTY GROUPS

	High Certainty Group	Low Certainty Group	Difference	P
Mean Frequency of Uncertain Statements	7.5	11.9	4.4	.14
Mean Frequency of Vague Modifying Statements	5.4	9.5	4.1	.22
Uncertain Statements and Vague Modifying Statements Combined	12.9	21.4	8.5	.09

TABLE IV
OUTCOMES OF TAT STORIES

	High Certainty Group	Low Certainty Group	Difference	P
Mean Frequency of Favorable Outcomes	6.5	5.8	0.7	.31
Mean Frequency of Unfavorable Outcomes	9.0	6.8	2.2	.05
Mean Frequency of Questionable Outcomes	4.5	7.3	2.8	.10

In summarizing the foregoing results, it seems that, on the whole, the two groups of subjects reveal no *striking* differences in their TAT stories. However, the data suggest that the high certainty group tells more stories in which their characters indicate needs for harm-avoidance, autonomy, and dominance, that their language is not quite so filled with value terms and cautious terminology, and that their stories have unfavorable outcomes more frequently than for the low certainty group.

Discussion

There are several possible interpretations for these results.

1. First, the criterion used in this experiment may not be a good one. Each of the certainty groups may be very heterogeneous groups. It is conceivable that several different sorts of individuals may give low certainty ratings in this light-buzzer situation, and that several different sorts of individuals may give high certainty ratings. There is no doubt that this criterion needs refinement. Perhaps a projective technique, or any other personality test, can predict more accurately if the criterion situation is a more specific one.

The adequacy of the certainty scale as a criterion would depend, to a great extent, upon its reliability and validity. A measure of the reliability was obtained by calculating the product moment coefficient of correlation between mean certainty scores on odd trials and even trials during the 120 trial series. The size of this correlation coefficient was .89. The validity of the certainty scale has not been determined. Previous investigations have shown significant degrees of correspondence between certainty of judgment and time required for making a judgment[2, 8], between confidence ratings and difficulty[3], and between level of confidence and amount of stimulus difference in making comparative judgments of two stimuli[4]. It would seem that these interrelationships furnish evidence of the validity of scales of certainty or confidence.[1]

In spite of the probable defects of the criterion, it is believed that the differences between the high certainty group and the low certainty group in this experiment are more than chance differences and that a more accurate assessment of their characteristics would reveal these differences.

2. Second, it is possible that certainty or confidence is not a general personality trait—that an individual may be certain in some situations and uncertain in others. If this is true, one could not expect even the best personality test to correlate with behavior in one specific situation. This problem of generality of certainty or confidence in behavior is itself a problem for investigation.

3. Third, there is of course the possibility that the TAT cannot predict this kind of behavior here called certainty. The writers believe that the TAT tells something about the individual, but in this case it doesn't reveal much about his certainty or confidence. In regard to this characteristic of behavior, it is questionable whether the small amount of suggestive information obtained is worth the effort of analyzing TAT stories.

It is possible that some of the cards of the TAT series are more suitable than others for eliciting stories in which certainty or confidence of the subject would be reflected. The writers' opinion is that many important items of information can be derived from TAT stories, if one only knows just how to analyze the data. The methods of analysis used in this investigation may not be the most appropriate possible. This problem—what is an appropriate analysis to make for a given purpose?—can perhaps be answered by a great many more trial investigations.

---

[1] There is some evidence for the validity of the scale used in the present investigation. When this scale was used in a learning experiment, it was found that learning curves based on certainty judgments closely paralleled learning curves based on frequency measures. Frequency, or percentage of responses, seem established as a measure of learning" page 283. Assuming that certainty or confidence increases as learning progresses, this measure of learning should furnish an adequate criterion for validation of the certainty scale. Similar evidence that certainty furnishes a valid subjective measure of learning is presented by Johnson and Rhodes[8].

SUMMARY

The problem of this experiment was to determine whether a characteristic of certainty or confidence can be assessed by means of the Thematic Apperception Test. Ninety subjects were asked to indicate their degree of certainty, on a five point scale, for their predictions of uncertain occurrences. More specifically, the subjects predicted, in response to various light combinations, whether a buzzer would follow. There was no basis for predictions of high certainty, inasmuch as the buzzer followed each light combination on fifty percent of the trials in random order. The ten subjects who gave the highest certainty ratings and the ten subjects who gave the lowest certainty ratings were administered the Thematic Apperception Test. Stories were analyzed by means of a Murray[6] need analysis and were also analyzed for language usage and for desirability of outcomes of stories. Few striking differences, but several suggestive differences, were found between the two groups of subjects. Various interpretations for these results were offered.

REFERENCES

1. BALKEN, EVA R., & MASSERMAN, J. H. The language of phantasy: III. The language of the phantasies of patients with conversion hysteria, anxiety state, and obsessive-compulsive neuroses. *J. Psychol.*, 1940, 10, 75-86.
2. FESTINGER, L. Studies in decision: II. An empirical test of a quantitative theory of decision. *J. exp. Psychol.*, 1943, 32, 411-423.
3. HERTZMAN, M. Confidence ratings as an index of difficulty. *J. exp. Psychol.*, 1937, 21, 113-119.
4. JOHNSON, D. M. Confidence and speed in the two-category judgment. *Arch. Psychol.*, 1939, 241, 1-52.
5. JOHNSON, D. M., & RHOADES, C. Measurement of a subjective aspect of learning. *J. exp. Psychol.*, 1941, 28, 90-92.
6. MURRAY, H. A. *Explorations in Personality.* New York: Oxford University Press, 1938.
7. POSTMAN, L., & EGAN, J. P. Experimental psychology: an introduction. New York: Harper & Bros., 1949.
8. VOLKMANN, J. The relation of the time of judgment to the certainty of judgment. *Psychol. Bull.*, 1934, 31, 672-673.

SZONDI RESULTS IN GROUP TESTING WITH COLLEGE STUDENTS

DAVID COLE AND EUGENE ROBERTS

Occidental College

PROBLEM

The purpose of this study was to investigate the Szondi Test performances of undergraduate college students, with the aim of obtaining a concrete picture of what trends may characterize such a population.

METHOD

The subjects of the study were fifty undergraduate students, enrolled in an introductory course in psychology. The testing was by a group technique, where slides of the six Szondi sets were projected on a screen in front of the group. In this group technique, each Szondi set is presented in sequence with standard Szondi instructions, and the subjects are asked to indicate their preferences and rejections on individual charts. The entire procedure requires about seven minutes for a given presentation, and the testing was continued at the start of each class hour for ten consecutive class meetings. These meetings covered three and one third weeks. Permission to produce the slides was obtained from the publisher of the Szondi materials, Hans Huber, Bern, Switzerland. The group technique appeared entirely satisfactory from the standpoint of mechanics of administration. In a study in progress, L. J. Borstelmann and W. G. Klopfer, of the University of California, have demonstrated that the results obtained through use of the group technique do not differ from those obtained via individual administration.

The results of the testing have been analyzed to reveal:

1. To what extent are the pictures of the test utilized or rejected on more than a chance basis?

2. Once a picture is selected, does the nature of the selection have a constant direction, or may the picture be randomly selected as "most preferred," and "least preferred?"

3. Are there certain pictures within a given clinical group that are more generally liked while others within the same clinical category are more often disliked, by our subjects as a whole? In her analysis of the Szondi record of Gregor[1], Deri implies that the stimulus values of individual pictures within a given group are not constant, but does not give figures supporting this.

Twenty three of the students used in this study were men and twenty seven were women. The median age of the men was twenty, with a range from eighteen to twenty nine. The median age of the women was nineteen, with a range from eighteen to twenty four. The subjects were predominately sophomores.

An effort was made to obtain ten records from each participant. Seventeen of the women and eleven of the men met this criterion. Only one woman had as few as seven sessions; among the remaining twelve men, ten completed nine, one eight, and one five sittings. There are a total of 213 records from the men students and 255 from the women.

Results

As Table I indicates, selection of a given Szondi picture does not occur on a chance basis in most cases. For the women, among the forty eight pictures, thirty five are either selected or ignored with a frequency which reaches either the .01 or .05 confidence level. Similar figures hold for the men.

Table II shows that given pictures are in most cases predominately liked or disliked. Most of the pictures have clear-cut affective value for our subjects.

Table I
Selection of Individual Szondi Cards By Set

Clinical Category	Selection Frequency Men	Women	Clinical Category	Selection Frequency Men	Women
Set I			Set IV		
H	33 % **	32 % **	H	63 % **	43 % *
S	45 %	38 % **	S	53 %	51 %
E	74 % **	90 % **	E	33 % **	37 % **
Hy	26 % **	13 % **	Hy	27 % **	44 % *
K	72 % **	71 % **	K	51 %	53 %
P	52 %	12 % **	P	54 %	52 %
D	59 % *	64 % **	D	78 % **	79 % **
M	40 % **	77 % **	M	40 % **	42 % *
Set II			Set V		
H	40 % **	39 % **	H	70 % **	72 % **
S	60 % **	53 %	S	17 % **	33 % **
E	54 %	48 %	E	59 % *	57 % *
Hy	90 % **	81 % **	Hy	45 %	42 % *
K	15 % **	22 % **	K	43 % *	52 %
P	40 % **	59 % *	P	61 % **	50 %
D	33 % **	48 %	D	36 % **	38 % **
M	68 % **	50 %	M	67 % **	55 %
Set III			Set VI		
H	45 %	35 % **	H	50 %	64 % **
S	55 %	36 % **	S	47 %	55 %
E	33 % **	30 % **	E	54 %	48 %
Hy	78 % **	71 % **	Hy	46 %	61 % **
K	65 % **	82 % **	K	69 % **	60 % **
P	22 % **	33 % **	P	40 % **	31 % **
D	20 % **	45 %	D	62 % **	44 % *
M	82 % **	68 % **	M	32 % **	35 % **

* Exceeds chance frequency (50 per cent) at .05 confidence level.
** Exceeds chance frequency (50 per cent) at .01 confidence level

TABLE II
PREFERENCE FOR INDIVIDUAL SZONDI CARDS BY CLINICAL CATEGORY

Clinical Category	Men		Women	
	Times Selected as Liked	Times Selected as Disliked	Times Selected as Liked	Times Selected as Disliked
Homosexual				
Set I	25	46 *	30	52 *
Set II	54	31 *	27	72 **
Set III	57	39	51	38
Set IV	53	81 *	42	68 *
Set V	103	46 **	148	35 **
Set VI	89	17 **	147	18 **
Sadist				
Set I	34	61 *	17	81 **
Set II	81	47 **	48	87 **
Set III	90	26 **	55	37
Set IV	68	46 *	96	34 **
Set V	25	11 *	63	21 **
Set VI	26	73 **	11	130 **
Epileptic				
Set I	150	7 **	217	13 **
Set II	28	87 **	23	99 **
Set III	54	16 **	66	11 **
Set IV	28	42	41	53
Set V	68	58	65	80
Set VI	2	112 **	10	113 **
Hysteric				
Set I	51	4 **	17	16
Set II	2	189 **	1	206 **
Set III	7	160 **	15	167 **
Set IV	21	37 *	74	38 **
Set V	20	75 **	24	83 **
Set VI	79	19 **	115	41 **
Catatonic				
Set I	0	154 **	3	178 **
Set II	18	13	44	12 **
Set III	3	136 **	2	207 **
Set IV	84	25 **	100	35 **
Set V	30	62 **	28	105 **
Set VI	119	29 **	104	50 **
Paranoid				
Set I	105	6 **	19	12
Set II	72	12 **	132	18 **
Set III	22	24	77	8 **
Set IV	111	4 **	129	4 **
Set V	106	23 **	86	42 **
Set VI	59	26 **	42	37
Depressive				
Set I	7	118 **	15	149 **
Set II	58	13 **	117	6 **
Set III	30	13 *	111	4 **
Set IV	19	148 **	1	200 **
Set V	7	70 **	29	69 **
Set VI	1	132 **	13	99 **
Manic				
Set I	55	30 *	191	6 **
Set II	118	26 **	119	9 **
Set III	159	16 **	132	42 **
Set IV	43	42	27	79 **
Set V	64	78 *	68	73
Set VI	54	15 **	69	20 **

Maximum number of selections for men: 213
Maximum number of selections for women: 255
* Like-dislike trend varies from chance at .05 confidence level.
** Like-dislike trend varies from chance at .01 confidence level.

For some of the pictures there are definite sex differences in terms of affective value. The sadist picture in set II is selected as "liked" by the men at the .01 level of confidence, and as "disliked" by the women at the same confidence level. Further, the predominant response to a given picture, once selected, changes direction for pictures within the same clinical category. For example, among the six catatonics, for the women the group trend is to select three of them as liked at the .01 level of confidence, while the other three are selected as disliked, to the same degree of statistical probability.

It would seem that this variance within a clinical category must be regarded as crucial both in interpretation of individual test results and in consideration of test rationale. While not involving directly the question as to whether or not the clinical categories used by the test display uniform facial characteristics, these data do suggest that other, possibly more superficial characteristics of the individual cards are of considerable importance in determining group trends. The presumed commonality of certain aspects of the pictures in a given clinical category do not appear to be crucial in determining our subject's reactions to the pictures.

Table III indicates the frequency of specific reactions to specific categories. Only a few of these show statistically reliable sex differences.

Table IV shows the frequency with which particular vector constellations occur. The chief sex differences lie not in the number of individuals who obtain a given vector pattern, but in the frequency with which persons of the given sex repeatedly give the pattern, once it occurs in their profile at all. It would appear that for interpretive purposes, the chief sex differences are in terms of the degree of consistency of particular test reactions, not in terms of occurrence versus non-occurrence.

Although it will not be dealt with extensively, readers who study Table IV will note that many of our results are directly in line with what might be predicted from Deri's *Introduction to the Szondi Test.* In the Sexual Vector for example, plus H proves predominately a feminine reaction and plus S chiefly masculine. Plus E, minus Hy is by far the most common pattern in the Paroxysmal Vector. The frequency of the anxiety indicative minus E, minus Hy

TABLE III

GROUP TRENDS IN SZONDI FACTOR SCORES

		Sexual Vector Men	Sexual Vector Women			Ego Vector Men	Ego Vector Women
Plus	H	39%	49% *	Plus	K	10%	05%
Minus	H	24%	21%	Minus	K	46%	56% *
Open	H	27%	21%	Open	K	20%	13%
Ambiv	H	10%	09%	Ambiv	K	24%	26%
Plus	S	31%	17% **	Plus	P	71%	62% *
Minus	S	27%	33%	Minus	P	03%	08%
Open	S	30%	35%	Open	P	23%	28%
Ambiv	S	12%	15%	Ambiv	P	03%	02%
		Paroxysmal Vector Men	Paroxysmal Vector Women			Contact Vector Men	Contact Vector Women
Plus	E	24%	41% **	Plus	D	04%	09%
Minus	E	21%	25%	Minus	D	66%	58%
Open	E	26%	17½% *	Open	D	20%	15%
Ambiv	E	29%	17½%	Ambiv	D	10%	18% *
Plus	Hy	02%	09%	Plus	M	59%	62%
Minus	Hy	70%	64%	Minus	M	10%	10%
Open	Hy	15%	15%	Open	M	15%	16%
Ambiv	Hy	13%	12%	Ambiv	M	16%	12%

* Percentage difference significant at .05 confidence level.
** Percentage difference significant at .01 confidence level.

TABLE IV
GROUP TRENDS IN SZONDI VECTOR PATTERNS

				Frequency of Pattern Occurence		Number of Subjects Showing Pattern	
Sexual Vector				Men	Women	Men	Women
Plus	H,	Plus	S	23	27	8	10
Plus	H,	Minus	S	24	39	7	13
Plus	H,	Open	S	17	46 **	7	9
Plus	H,	Ambiv	S	19	13	7	6
Minus	H,	Plus	S	13	13	6	7
Minus	H,	Minus	S	13	17	5	6
Minus	H,	Open	S	25	20	9	5
Minus	H,	Ambiv	S	1	3	1	1
Open	H,	Plus	S	22	8 **	6	5
Open	H,	Minus	S	16	21	6	10
Open	H,	Open	S	17	11	7	6
Open	H,	Ambiv	S	2	12	2	6
Ambiv	H,	Plus	S	9	5	5	3
Ambiv	H,	Minus	S	4	7	3	3
Ambiv	H,	Open	S	5	13	5	6
Ambiv	H,	Ambiv	S	3	0	1	0
Paroxysmal Vector							
Plus	E,	Plus	Hy	2	10	2	7
Plus	E,	Minus	Hy	35	63 *	16	17
Plus	E,	Open	Hy	5	21	4	9
Plus	E,	Ambiv	Hy	10	11	7	7
Minus	E,	Plus	Hy	1	5	1	5
Minus	E,	Minus	Hy	36	45	12	15
Minus	E,	Open	Hy	3	6	3	5
Minus	E,	Ambiv	Hy	5	7	5	2
Open	E,	Plus	Hy	0	1	0	1
Open	E,	Minus	Hy	39	32	13	15
Open	E,	Open	Hy	8	5	3	4
Open	E,	Ambiv	Hy	8	6	4	3
Ambiv	E,	Plus	Hy	1	6 *	1	4
Ambiv	E,	Minus	Hy	39	23 **	13	13
Ambiv	E,	Open	Hy	15	5 *	6	4
Ambiv	E,	Ambiv	Hy	6	10	3	6
Ego Vector							
Plus	K,	Plus	P	19	8 **	5	4
Plus	K,	Minus	P	1	0	1	0
Plus	K,	Open	P	3	5	2	3
Plus	K,	Ambiv	P	0	0	0	0
Minus	K,	Plus	P	65	91	20	22
Minus	K,	Minus	P	2	17 **	1	6
Minus	K,	Open	P	27	32	11	14
Minus	K,	Ambiv	P	4	3	3	2
Open	K,	Plus	P	29	22	10	7
Open	K,	Minus	P	3	1	2	1
Open	K,	Open	P	7	7	4	6
Open	K,	Ambiv	P	1	2	1	2
Ambiv	K,	Plus	P	38	40	11	11
Ambiv	K,	Minus	P	0	1	0	1
Ambiv	K,	Open	P	13	26 *	7	8
Ambiv	K,	Ambiv	P	1	0	1	0

TABLE IV (CONCLUDED)
GROUP TRENDS IN SZONDI VECTOR PATTERNS

Contact Vector				Frequency of Pattern Occurence		Number of Subjects Showing Pattern	
				Men	Women	Men	Women
Plus	D,	Plus	M	5	10	3	5
Plus	D,	Minus	M	1	3	1	3
Plus	D,	Open	M	1	7 **	1	4
Plus	D,	Ambiv	M	2	3	1	2
Minus	D,	Plus	M	90	100	16	22
Minus	D,	Minus	M	10	17	9	8
Minus	D,	Open	M	26	17 *	8	10
Minus	D,	Ambiv	M	15	14	7	10
Open	D,	Plus	M	21	19	7	9
Open	D,	Minus	M	9	5	3	3
Open	D,	Open	M	2	1	1	1
Open	D,	Ambiv	M	9	13	4	7
Ambiv	D,	Plus	M	8	29 **	4	10
Ambiv	D,	Minus	M	2	2	2	3
Ambiv	D,	Open	M	3	15	3	6
Ambiv	D,	Ambiv	M	9	0 **	3	0

* Differences significant at .05 confidence level.
** Differences significant at .01 confidence level.

combination appears with the expected frequency for our adolescent group. The high incidence of plus P among college students, which she predicts, is demonstrated.

Of particular interest in the Ego Vector is comparison of our results with the stages of ego development outlined by Deri. Readers who wish to compare our findings with her material will note that our population appears at the adolescent stage of development. This slight measure of immaturity might be anticipated from the general socio-economic level of the participants.

CONCLUSIONS

1. Group administration of the Szondi test is feasible and highly promising in expediting research.

2. Selection of and reaction to, various Szondi pictures show consistent trends for groups of similar subjects. The pictures appear to differ markedly in affective tone. The significance of selection or rejection of given pictures in individual cases should be considered in light of these findings. As a hypothetical example, a female subject in our group who obtained a minus K by selecting as "disliked" the two catatonics so overwhelmingly accepted by the group as a whole would appear to deviate from the group markedly, although her minus K was typical of the group reaction to the catatonic category.

3. Different pictures within the same clinical group have been shown to produce decidedly different group reactions. This finding seems to throw some question on the extent to which facial characteristics common to a given diagnostic group are of chief importance in determining Szondi choices. This finding, in view of the reported clinical success with the test, makes the search for an adequate rationale more demanding.

4. Many of the statements of Deri have been borne out. The extent to which one may make accurate predictive statements concerning group trends in test responses is of importance in determining the ultimate diagnostic utility of the test instrument. The fact that such statements may be made concerning reactions to the Szondi Test stimuli increases the potentialities of the test.

BIBLIOGRAPHY

1. DERI, SUSAN K., "The Szondi Test," in "The Case of Gregor: Interpretation of Test Data," *Rorschach Research Exchange* and Journal of Projective Techniques, 32, 1949, 438-443.
2. DERI, SUSAN K., *Introduction to the Szondi Test.* New York: Grune & Stratton, Inc., 1949. 354 pp.

THE TEST-SITUATION AND THE PROBLEM OF PREDICTION

SEYMOUR B. SARASON

Yale University

I

The purpose of the present paper is to state what seem to be the important determinants of behavior in a testing situation and to examine their implications for the problem of prediction. It should be made very clear at the outset that there is nothing new in the statements of these determinants; nevertheless, they require listing and brief description because the implications to be drawn from them seem to call for a revised approach to the problem of prediction. Attention will be given below to those determinants which are present regardless of the sex, age, and social class of the individual. In addition, there is no intention in the following discussion to convey the impression that the determinants are independent of each other; on the contrary, it is implicitly assumed that they are interdependent variables.

A. *The nature of the stimulus materials.* The content of the individual's responses, the accompanying expressive behavior, the needs and attitudinal factors engendered are in part a function of the nature of the stimulus materials. How an individual will react to a thematic card cannot be determined from knowing how he reacted to a comprehension item on the Binet.[1] Regardless of the continua (familiarity, physical attributes, shape, etc.) along which one may order stimulus materials, the behavior of the individual will in part be a function of that point on a particular continuum that the stimulus material falls.

B. *The nature of the instructions.* An individual who is told that he may respond in whatever way he wants, or that he must complete a task within a certain time limit, or that he must reproduce a stimulus configuration in an exact manner will not necessarily react in a similar fashion to these different instructions.

C. *The purpose of the testing.* How the individual behaves in the the testing situation depends on (a) what reasons the psychologist gives for the testing and (b) what the individual has concluded, either before or during testing, are the reasons he is in the particular situation. For example, given three individuals who are alike in several respects, if the first feels that whatever the tests show will "be held against him", he will respond differently from the second who feels that the testing situation is fun, or from the third who looks upon it as downright silliness. The opinion might be ventured here that more than a few psychologists assume that whatever reasons for the testing they have given to the subject are believed and accepted.

D. *The time and place of testing.* An individual's behavior is a function of when he is being tested and how many times he has experienced the testing situation. What may be equally important is where the individual is being tested: a general hospital, a mental hospital, a school, a psychology laboratory, a private office, or some kind of assessment center.

E. *The psychologist.* The behavior of the individual in the testing situation is in part a function of the sex, age, and status of the psychologist. If the purpose of an experiment is either to measure the degree of suggestibility of a subject or to inculcate a feeling of failure, it cannot be assumed that the behavior of the subject will be the same regardless of whether the psychologist is a graduate assistant or a professor in the department.

F. *Attitudinal factors related to previous conditions of learning.* The testing situation engenders varying attitudes toward self and others in different individuals. In some cases feelings of inadequacy,

[1] In the subsequent discussion words like behavior, reactions, or responses refer not only to what an individual says or does but also to the accompanying needs and attitudes which may be overtly manifested or inferred from other behavioral data.

anticipations of failure, and need to avoid anticipated ridicule are aroused even before testing begins. Whatever the needs, attitudes and defenses called forth by the situation, they are a function of previous learning experiences. What one observes in the testing situation is a restricted sample of the ongoing learning process—the content of the sample (i.e., the observed behavior) being a function of past situations which are in some ways similar to the present one. This similarity may be due to several factors among which are the specific stimulus materials and the strength of the previous internal responses. For example, some children whose difficulty in reading is a function of emotional factors may respond less emotionally to the learning of arithmetic; other children may respond with equal involvement to both situations. In the former instance it is as if some kind of discrimination has been made between the two situations, so that the responses elicited by them appear to be rather different; in the latter it is as if a very strong emotional response to one situation has generalized to what objectively appear to be dissimilar stimulus configurations.

II

For the purposes of the present discussion the problem of prediction may be put as follows: from what and to what is one predicting? A common practice is to assign to an individual a score based on his formal test responses and then to predict either his score on another test or his mode of response to another situation. There are several assumptions which usually are made, either explicitly or by inference, when such a procedure is followed, but seldom is there any attempt to evaluate the role and applicability of these assumptions. For example, while the subject's score on the initial test necessarily reflects the variables described in Section I, there is rarely any consideration given to the extent to which each of these variables is either involved or unrelated. If the latter it may be disregarded; if the former it must be taken into consideration if the predictions made are to be based on pertinent variables and not contaminated by the irrelevant ones. A second assumption is that the second situation, regardless of surface similarity or dissimilarity to the first, elicits the same needs and attitudes as did the first situation, a process of generalization of response being facilitated. Since in most studies one does not know or take into account the nature and strength of the variables in the first situation, there is usually little or no evidence for making this second assumption. In those instances where the investigator wishes to predict from one kind of stimulus situation to an obviously different one it is assumed that either the individual will get similar scores because the same kind of behavior is engendered despite the objective differences in the stimulus situations, or that scores will be negatively correlated because of the differences in the stimulus materials. More often than not the scores used in such studies do not reflect the determinants of the behavior but the end product of their interaction.

Three things generally happen to such studies: the results are disappointing and are not reported in the literature; the results are disappointing and are published with the blame directed at the inadequacies of the test; the results are encouraging, which usually means that the correlations are significant above zero, most of the variance is unaccounted for, and prediction to the next case one involves a hazardous procedure.

The point that deserves emphasis is that studies in the prediction of behavior on the basis of psychological tests (or any interpersonal situation) have proved disappointing. It is the contention of the writer that the disappointing results are largely due to the failure to take into account commonly acknowledged significant determinants of behavior, lack of information about the relative importance of these determinants, and the absence of procedures for their measurement. Without knowledge of these determinants one is simply unable to answer the question: from what and to what is one predicting? In the following sections several studies will be discussed in order to concretize some of the points made.

III

Several years ago a study was begun with the aim of discovering how best to predict who will make a good or bad clinical psychologist.[2] Students who had been accepted for graduate work in clinical psychology were assembled in various assessment centers for several days, a wide variety of tests administered to them, "stress interviews" conducted, observations in different kinds of OSS-type of situations made, autobiographies obtained, etc. After these students had been in graduate school for several years they were rated by the staff of the university departments who had contacts with them, and also by the supervisory staff in the field agencies. The predictions made at the time of original assessment proved to be disappointingly poor.

The question, of course, is why such poor predictions? There are undoubtedly many factors (unreliability of ratings, for example) which may have contributed to the findings, but attention here will be focussed on the points discussed earlier in the paper.

The first comment that could be made concerns the students' reaction to the stated purpose of the assessment procedure. Although each student knew that he had been accepted for graduate school and had been told that the results of the assessment would in no way affect him, it does not follow that all participants became "ego-involved" to the same degree. Some undoubtedly came to the situation with heightened anxiety and some with little or none. In connection with this it should be pointed out that the assessors were "names" in psychology and experts in the particular procedure they employed. It may be assumed that there was wide variation in the degree and nature of the motivation of the students. In short, individuals probably differed not only in the specific drives or needs engendered but in the degree to which any of these drives was involved. If the kind and level of a drive aroused by a specific stimulus is one of the determinants of the overt response, it is theoretically possible and probable that classification only on the basis of the overt response would be misleading and incomplete. Because two individuals give similar responses to a certain test or respond in an overtly similar manner to a stress situation does not allow one to ascribe their performance to similar psychological factors.

The relevance of several of the other determinants of behavior for the assessment study can be inferred from an examination of the assumption that the various stimulus situations and the ensuing behavior are representative of the situations these students would encounter in graduate school and in field placement. Because a student "went to pieces" in some of the procedures does not mean that attending classes, doing research, administering tests, or doing psychotherapy will evoke similar inadequate behavior. Conversely, the student who never appeared to be perturbed during assessment (perhaps because he knew his future status was in no way at stake?) is not necessarily the one who will sail smoothly through his academic and clinical duties. Aside from the dubious relevance of the criteria, the point should be reiterated that it is only when one knows the "internal" significance of a situation for the individual that one can describe and predict the situations which would elicit similar behavior. Because two children lose their parents at whatever is considered a crucial stage of development obviously does not allow one to say that they were similarly affected by the deprivation. It is not enough to know either the nature of the external stimulus or the content of the individual's formal response. One must also know the kinds of internal behavior (attitudes toward self or others, set, level of anxiety, fantasy content, etc.) with which the overt behavior seems related. It is as important to know the nature of internal as well as external stimuli. Research interest for

[2] This study has not yet appeared in the literature. The present comments are based on a very brief presentation of preliminary findings made by Dr. Lowell Kelley at the Boulder Conference on Graduate Education in Clinical Psychology. What critical comments the writer makes here should not obscure the singular, and perhaps historical, importance of this study in forcing clinical psychology to look very critically at some of its commonly accepted procedures.

long has been focussed on the external stimulus and overt behavior with too little attention to devising techniques for getting at and evaluating internal stimuli, studying their relevance for any classification of subjects, and investigating how such data can be reflected in an actuarial statement.

The significance of the assessment study is that it is representative of current thought and practice in clinical psychology. It illustrates how clinical psychologists approach the problem of prediction and how inadequate the approach is. But it would be unfortunate if the major implication drawn from the study is that our present tests are inadequate. Such an implication is in part a justifiable one but to emphasize it would have the result of overlooking our neglect of the determinants of test-situation behavior. In these days of the "total personality", current practice seems inadequate to its assessment.

IV

A concrete clinical example may clarify the previous discussion. There are certain children whose pattern of behavior is sufficiently similar so as to warrant a common label: primary behavior disorder. An abnormal amount of aggressiveness, inability to form satisfactory relationships with people or objects, apparent absence of guilt feelings —these are the characteristics, present since early childhood, which apply to these children and are considered a patterned reaction to unfavorable environmental influences.

A not infrequent experience of therapists with such children is the marked discrepancy between the case history description of behavior and the child's behavior during the initial contacts.

"It has been observed frequently that a child with a severe primary behavior disorder is friendliness and sweetness personified during the first few interviews, but thereafter, without fail, returns to the old aggressiveness. Even one of our most experienced and skilled workers felt a question arise in his mind about the correctness of the diagnosis when the boy he was interviewing for the first time was so nice that it seemed impossible to him for this to be the same little devil described in the intake worker's report. However, he soon found out . . . ."[2]

It invariably happens that when a child with conduct disorders, a child upon whom all sorts of threats and forms of punishment have been tried unsuccessfully, is finally referred to our agency, the parents have told him that we are going to show him that he cannot go on as he has done, that we are going to be very strict and severe, and so on. Naturally, the child expects unfriendliness, harshness, and threats. To his great surprise, he finds a worker who is quiet and friendly, who is interested in every little thing he mentions. She speaks to him in a civil way and does not scold him. This attitude takes him by surprise and he finds himself disarmed, having no reason or justification for a display of his usual aggressiveness. His hostility finds itself in a vacuum with nothing to fight and the surprising attitude of the worker throws him back into an attitude of friendliness. We have been using the term "surprise reaction" to indicate this behavior: finding to his surprise that there is nothing provoking his hostility, the child reverts, as it were, to the kind of behavior he may have shown before his disorder started. We would be inclined to expect that after the child has had the experience of a pleasant and warm relationship, he might be eager to continue in it, thankful to have found at last a response that he had not met with in a long time, either in his home or in school. But we find ourselves in error very soon and we want to understand what it is that brings about this change."

It would be expected that if one were to predict the behavior of these children on the basis of initial contacts with them, the accuracy of one's predictions would probably be very poor. Unless one viewed the observed behavior in terms of the determinants discussed earlier in this paper, then the child's reactions, present and future, are inexplicable.

V

Lantz's[1] study of "some dynamic aspects of success and failure", while not one specifically concerned with the

problem of prediction, illustrates some of the points that have been discussed. The design of the study is briefly given below:

1. 212 nine-year-old boys were studied.
2. They were given 9 items from the Revised Stanford-Binet Scale (L).
3. Each child was put in a ball game situation in which the task was to remove a ball from a box using only one hand. "The ball is to be secured from the box three times in three different ways. Two experiences of success precede the third experience of success or failure and the chosen prize is given only for success on this third problem." Success or failure on the third problem was a function of the subject and not the examiner.
4. The ball game tasks were immediately followed by a comparable set of test items from the Form M scale.

Without going into the details of the analysis it will suffice for the present discussion to state Lantz's conclusion that "The experiences of success and failure do influence some mental processes more significantly than others. The depressing effect of the experience of failure in decreasing scores is greater than is the exhilarating effect of success in increasing them."

Since success or failure in the third ball game task was not predetermined by the experimenter but by the subject's performance, the question arises: why did some boys succeed and others fail in the third task? All the boys appeared to be highly motivated to succeed in order to win a desired prize. The possibility must be considered that failure was in some way a function of the kinds of responses engendered by this stressful situation. If such a possibility has merit, then one must ask if the factors making for a failure performance were also present in the initial test situation (Form L). The child whose reactions to the ball game situation had a disorganizing effect on his performance may also have been the one whose initial Binet performance was also affected by these covert responses. Therefore, it does not seem justified to assume that performance on the second Binet testing was being influenced solely by the experience in the ball game situation. In fact, although Lantz maintains that there was no significant difference between her success and failure groups in the initial Binet testing, the critical ratio between the means of 2.63 would indicate otherwise. Even if there was no difference between the two groups in initial testing, one cannot assume that similar scores stand in the same relationship to the determinants of behavior in the Binet situation. If one had a knowledge of these determinants it might well be that in the case of two individuals with identical scores the prediction might be made that one would fail in the ball game situation while the other would not.

VI

In his studies of adult intelligence Wechsler[3] found a rather steady decrease in intelligence test scores beginning at approximately twenty-five years of age. Wechsler concluded that "the decline of mental ability with age is part of the general senescent process of the organism as a whole." This line of reasoning was supported by the similarity between the curve of mental decline and those of vital capacity and brain weight. While one cannot doubt the idea of a normal senescent process, it does not follow that the difference in test scores between groups of 25 and 50 year old men are solely due to such a process. When two such groups are viewed in light of what have been called the determinants of test behavior, it seems reasonable to assume that the kinds of overt and internal behavior aroused by the test situation are not necessarily the same for the two groups. In a man of 50 the fear of appearing inadequate on an intelligence test might be inordinately strong not only because he *has* slowed up but because he is concerned that he *may* have slowed up. Middle-aged people undoubtedly differ considerably in their need to feel that they are as alert mentally as ever. With a man of 50 many years may have elapsed since he last was in a test situation. The 50 year old man being tested by a much younger person is probably

not in the same situation as a 25 year old man being tested by a somewhat older person. In short, one does not know the extent to which differences in scores between such age groups are a function of behavioral factors varying with but not caused by age. Again it would seem that interpretations of findings based on the end product of a highly complicated behavior sequence (anticipatory reactions based on previous learning ⟶ external stimulus ⟶ internal reactions and drives ⟶ articulated response) are theoretically likely to be very incomplete.

VII

The chief conclusion to be drawn from the above considerations is that research in the problem of prediction has neglected important determinants of behavior. One of the reasons for this neglect is probably the failure to view the problem in terms of a systematic theory of behavior. An explicitly stated theory serves at least three very practical purposes: (a) it states publicly what one considers to be the important variables and points the way for experimental verification; (b) it makes the nature of one's disagreements with other researchers more clear and makes resolution of differences by crucial experiments more likely; (c) it makes the inadequacies of the shot-gun approach more clear and makes the training of newer generations of students not only easier for the teacher but far more stimulating for the student. Until there are attempts on the level of theory to identify the determinants of behavior, to state the principles by means of which they become related and how they can be measured, putting one's eggs in the basket of "conventional clinical procedures" is not likely to pay off.

Another reason for the neglect of some of the important determinants of behavior is that techniques—observational, physiological, biographical—for evaluating the internal responses to stimulus configurations are either not available or crudely developed. What the subject does not verbalize to the observer (the extent, content, and relevance of his self-verbalizations before and during presentation of the task) is an intervening process which, despite its apparent importance, is very difficult to assess and so has been little studied. Any attempt at theory building must take account of these intervening processes and give impetus to the devising of techniques for their evaluation.

A concluding note: it might be argued by some that the uniqueness of the individual precludes the possibility of predicting his behavior from a knowledge of the behavior of others. In order to predict, the argument runs, one must study *that* individual. It is not the purpose of this paper to examine the implications of this argument for the possibility of establishing a science of behavior. This much might be said in partial answer: if in studying an individual case the psychologist does not start out with a systematic formulation of what are the determinants of behavior, he automatically restricts the kinds of things he will observe, and does not possess a basis for utilizing his wrong predictions as a means of sharpening his handling of the next case he meets. There would be nothing more comforting than to approach the individual case with the knowledge that one's activities have a sound basis in systematic theory and experimental evidence. It may be one would not know what specific weights to give to variables, but at least one would know what the variables are—a situation which does not obtain when we study the individual case today.

Bibliography

1. Lantz, B. Some dynamic aspects of success and failure. *Psychol. Monogr.*, 1945, *59*, No. 1.
2. (Staff: Jewish Board of Guardians) *Primary behavior disorder in children.* New York: Family Welfare Association of America, 1945.
3. Wechsler, D. *Measurement of adult intelligence.* Baltimore: Williams and Wilkins, 1941 (2nd ed.)

MENTAL AGE SCORES FOR THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN

WILLIAM R. GROVE

Phoenix (Arizona) Elementary Schools

Introduction

An event of considerable importance to psychologists charged with serving children in life situations is the recent publication of the Wechsler Intelligence Scale for Children (WISC)[7]. It ranks with the contribution of Arthur[1] and is a step beyond Arthur in two important respects. First, it is technically more adequate than the Arthur scale in statistical treatment. Second, it bridges the verbal-nonverbal gap with comparable research data not adequately provided by Arthur. In using the Arthur scale comparisons between verbal and nonverbal abilities remain to some extent a matter of clinical guesstimate which can never be very satisfactory. Another practical advantage is that the WISC requires less time for administration than the Arthur scale without sacrificing reasonable reliability. Moreover, the test materials of the WISC are less cumbersome and bulky than the Arthur equipment, and also less expensive.

The important question of the basic validity of the two scales is less clear. Arthur relies for evidence of validity principally upon correlations with the Stanford-Binet[3]. Wechsler's validity data applies principally to his original scale[5] which was standardized on adults. The WISC is essentially a restandardization of Form II of the Wechsler-Bellevue Scale[6] with certain extensions and revisions of administrative technique designed especially for younger age levels and for lower ranges of ability. Basic research data on the validity of the WISC itself are urgently needed, however.

In the opinion of the writer the WISC has one serious weakness in the fact that Wechsler has abandoned the mental age concept without providing an entirely satisfactory substitute for it. His reasons for abandoning mental age scores are given in detail in the test manual[7] *p. 2f.* in an excellent statement of the case against the MA. But Wechsler either misses or ignores the point that mental age scores serve one useful purpose in the process of test interpretation for which he has not provided a suitable substitute. Wechsler alludes to this problem when he says:

> "An MA is a definable level of test performance. If no further claims for it were made, the concept could be readily accepted as a practical method of defining levels of test performance."

With his complete renunciation of the mental age concept, Wechsler has also discarded any simple method of defining the level of test performance except in relation to the performance of other children of the same chronological age as that of the subject tested. It is useful to compare children with their age mates, but it is also useful in certain situations to be able to make comparisons of a particular child's test performance with children older or younger than himself. In the case of the WISC this might have been easily accomplished had Wechsler provided us with a series of overlapping distributions for different age levels on the scale as a whole with some valid method of comparing a child's score on one distribution with comparable scores on other distributions. Unfortunately from our standpoint Wechsler's aversion to MA's was such that he did not even provide us with data of this sort for the WISC in readily accesible form.

Computation of MA Scores

However, although it is somewhat awkward and time consuming to do so, it is possible to make comparisons of test performances from one age level to another on the WISC. It is also possible to calculate MA scores for the WISC quite accurately and without making any of the faulty assumptions implicit in extrapolated MA scores.

In general, MA scores are most useful between the MA levels from 4 to 12. Above that level, increments of intellectual growth diminish rapidly in succeeding years and the MA concept tends to

become less useful for interpretative purposes. Actually it is possible to calculate accurate MA scores on the WISC for all levels from MA level 5 to MA level 15. In practice I never calculate MA's above 12 on this test. I find it sometimes helpful to estimate MA scores below 5 for certain comparative purposes, but the accuracy of such estimates cannot be checked because Wechsler did not standardize the test below that level.

The possibility of calculating an MA score for any WISC test performance that falls between the average test performance of five-year-olds and that of fifteen-year-olds arises from the fact that Wechsler has assigned a weighted score of 10 in his scoring tables to represent the average performance on each subtest of any particular chronological age group in his standardization population. It thus follows that on ten subtests the total weighted score for any particular chronological age group would be 100. Similarly on five subtests the average total weighted score would be 50, etc.

Since an MA score is defined as that test performance which is average for a particular chronological age group, we may calculate an accurate MA score for a child's WISC test performance by finding the chronological age group among Wechsler's scoring tables that will yield a total weighted score for the ten subtests of exactly 100. Thus if the WISC test performance of an eleven year old child when rescored on the WISC scoring table intended for children aged from 8 years 4 months through 8 years 7 months yields a total weighted score of exactly 100, it would be correct and accurate to say that this child had an MA of 8 years 6 months on the WISC (the midpoint of the chronological age group for whom that scoring table was intended). Such information about the child is quite useful as a fairly meaningful definition of the approximate comprehension level at which we might expect this child to be capable of functioning in school and in other life situations.

Unfortunately, the calculation of MA scores for most children on the WISC is not quite so simple as the example just mentioned. In most cases a scoring table cannot be found which will yield a total weighted score of exactly 100. In these cases two adjacent scoring tables must be used. One will yield a total weighted score that is slightly more than 100, while the scoring table just above it (i.e. for an older age group) will yield a total weighted score of less than 100. The correct MA can then be found by simple linear interpolation.

The actual procedure for calculating an MA score for the WISC starts by locating among the scaled scoring tables given by Wechsler those scoring tables which yield total weighted scores nearest to 100 for that particular child. There are two possible ways of doing this. One is by simple inspection. By this method the psychometrist looks for that scoring table which yields scaled scores closest to 10 for all subtests. He then tries out that scoring table and, depending on his success, moves to a higher or lower scoring table as his results may indicate. If the total weighted score obtained is more than 100, he must select a higher age group for trial; if less than 100, he must move to a lower age group.

The writer has found another procedure more expeditious. This is to obtain an estimated MA by treating the obtained Wechsler deviation IQ as though it were a true IQ. In this way the psychometrist multiplies the child's chronological age by the IQ used as a decimal fraction. The resulting estimated MA is not a true MA value, but may be used to locate the approximate scaled scoring table to be used in the calculation of a true MA value in the manner explained above.

An actual example of the calculation of an MA score is provided in Table I. The child's chronological age is 9-0, by Wechsler's instructions[7], *p. 23*, or 108 months. The Wechsler full scale deviation IQ is 86. .86 of 108 months are 93 months, or 7 years 9 months. This is the *estimated* MA.

This estimated MA tells us only where to start in the process of calculating an accurate MA. We now score the child's performance on the scoring table intended for children aged from 7-8 through 7-11. This yields a total weighted score of 101. It is thus necessary to

TABLE I

CALCULATION OF AN MA SCORE FOR THE WISC
ILLUSTRATIVE CASE X: TESTED 1–11–50; BORN 12–21–40
CA: 9–0–20 — WECHSLER FULL SCALE IQ: 86

VERBAL TESTS	Raw Score	Scaled Score 7–8 thru 7–11	Scaled Score 8–0 thru 8–3	PERFORMANCE TESTS	Raw Score	Scaled Score 7–8 thru 7–11	Scaled Score 8–0 thru 8–3
Information	6	6	5	Picture Completion	12	15	15
Comprehension	12	14	14	Picture Arrangement	22	11	11
Arithmetic	5	8	7	Block Design	6	9	9
Similarities	7	11	11	Object Assembly	15	10	9
Vocabulary	23	10	10	Coding B	20		7
Sum of Verbal Tests		49	47	Coding A	27	7	
				Sum of Performance Tests		52	51
Sum of all Tests		101	98				

MA (Lower Limit Midpoint	7–10	101 Scaled Score
MA (Upper Limit Midpoint)	8–2	98 " "
MA (Interpolated)	7–11	100 Scaled Score

rescore on the scoring table for the next higher age group. This yields a total weighted score of 98. By interpolating between the midpoints of these two scoring tables the true MA of the child is found to be 7 years and 11 months. This is an actual MA score and it is as accurate and as valid as Wechsler's standardization of the test makes possible.

This scoring sample also illustrates a further complication that occurs as a result of Wechsler's inclusion of two coding subtests to be used with different age groups. Thus Coding A is used with all children younger than eight years, and Coding B with all children eight years old and older. Scoring tables for chronological age groups below eight are for Coding A and above that for Coding B. In calculating MA scores the psychometrist must keep this in mind.

In testing children for whom I intend to calculate an MA score I always follow the practice of giving the coding subtest appropriate for that child's chronological age first in the standard manner. At a later time in the event that I will need the other subtest score for calculating the MA, I administer the other coding subtest also.

Sometimes it is found convenient to calculate an MA score for a child who was given only one coding subtest but it develops that the comparable score on the other coding subtest is needed. In these cases I use Table II. Although this conversion table may not be entirely accurate in all situations, it is probably accurate enough for all practical purposes. It was prepared by comparing adjacent scoring tables in the WISC manual.

TABLE II

CONVERSION TABLE FOR TRANSMUTING RAW SCORES ON ONE CODING SUBTEST TO RAW SCORES ON THE OTHER

Equivalent Raw Scores		Equivalent Raw Scores		Equivalent Raw Scores	
Coding A	Coding B	Coding A	Coding B	Coding A	Coding B
0–7	0	20	14	38	27
8	1	21	15	39–40	28
9	2	22	16	41	29
10	3	23	17	42	30
11	4	24	18	43	31
12	5	25	19	44	32
13	6–7	26–27	20	45	33–35
14	8	28–29	21	46	36–37
15	9	30–31	22	47	38–39
16	10	32	23	48	40–41
17	11	33	24	49	42–44
18	12	34–35	25	50	45–93
19	13	36–37	26		

Interpolation of WISC Deviation IQ's

In the course of making regular clinical use of the WISC for more than six months I have found that there are times when a child's test performance should be scored on two adjacent scoring tables and his IQ found by interpolation. These are cases where the child's chronological age falls appreciably above or below the midpoint of the scoring table used in his case. If the child's chronological age is appreciably below the midpoint of the scoring table his IQ may be too low. If his chronological age is appreciably above the midpoint, of the scoring table his IQ may be too high.

In some cases the error resulting from failure to interpolate is negligible, in others it is of serious practical concern. Our study has revealed no simple method of estimating the size of the error in any particular case without actual tryout. In general, errors of this sort tend to be negligible at and above a CA of 13 years. Below that level discrepancies of as much as seven points have been found between IQ's calculated by using chronologically adjacent scoring tables. When it is considered that a difference of a few days in the date of testing can determine which of two adjacent scoring tables shall be used, it becomes at once apparent that for accurate clinical work a discrepancy of almost one-half of a standard deviation in final score is too great.

There is no more dramatic way to illustrate this problem than by presenting the scoring of an hypothetical test record in detail in Table III. This case was actually tested by the writer, but the birthdate was changed by a month or two in order to throw the problem into sharp relief. Obviously situations as extreme as this will seldom arise, but it is a matter of some concern to the clinician that they are possible. It is the possibility of serious error that we are stressing. The clinician must, by the nature of his task, be concerned with individual cases.

It will be noted from Table III that hypothetical Case A was tested on the day after his tenth birthday. It will be observed that this boy's full scale Wechsler IQ when calculated strictly according to rule is 100. However, had he been tested two days earlier, his calculated IQ would have been 107. This is a difference in percentile rank of from 50 in the former instance to 68 in the latter. Certainly two days difference in the date of testing should not lead to a discrepancy in final score of such magnitude. The interpolated value of 103 in this case is both practically and theoretically more representative of this boy's actual ability. We have admittedly selected an extreme instance to emphasize the *possible* errors that can occur by use of the present scoring tables as recommended by Wechsler.

Table III
WISC Scoring Samples

Hypothetical Case A: Tested 8–10–49 Born 8–9–39 CA: 10–0–1

Verbal tests	Raw Score	Scaled Score 10–0 thru 10–3	Scaled Score 9–8 thru 9–11	Performance tests	Raw Score	Scaled Score 10–0 thru 10–3	Scaled Score 9–8 thru 9–11
Information	13	11	11	Picture Completion	13	13	14
Comprehension	9	7	8	Picture Arrangement	26	9	10
Arithmetic	10	12	12	Block Design	15	9	10
Similarities	6	7	8	Object Assembly	20	9	10
Vocabulary	34	11	12	Coding	41	12	13
(Digit Span)	10	11	12	(Mazes)	15	9	10
Sum of Verbal Tests		59	63	Sum of Performance Tests		61	67
Prorated Scaled Score		49	53	Prorated Scaled Score		51	56
Verbal IQ		99	104	Performance IQ		101	108
Total Scaled Score		100	109	Full Scale IQ		100	107

Summary

This study has attempted to document the following three points:

1. Wechsler has provided clinical psychologists with another, valuable psychometric instrument by the downward extension and restandardization for children of the Wechsler-Bellevue Scale, Form II. The new test is called the Wechsler Intelligence Scale for Children.
2. In discarding the MA concept Wechsler has discarded more than he has provided an adequate substitute for. Nevertheless, MA scores can be accurately and validly calculated for the WISC between the 5 year and the 15 year levels.
3. Because of the coarseness of scaling in certain cases the careful psychometrist must sometimes interpolate between adjacent scoring tables in calculating Wechsler deviation IQ's for this test.

References

1. Arthur, Grace. *A Point Scale of Performance Tests: Vol. I, Clinical Manual.* New York: Commonwealth Fund, 1930.
2. Arthur, Grace. *A Point of Performance Tests: Vol. II, The Process of Standardization.* New York: Commonwealth Fund, 1933.
3. Terman, Lewis M. *The Measurement of Intelligence.* Boston: Houghton Mifflin, 1916.
4. Terman, Lewis M. and Merrill, Maud A. *Measuring Intelligence.* Boston: Houghton Mifflin, 1937.
5. Wechsler, David. *The Measurement of Adult Intelligence.* Baltimore: Williams & Wilkins, 1939.
6. Wechsler, David. The Wechsler-Bellevue Intelligence Scale, Form II: *Manual.* New York: Psychological Corp., 1946.
7. Wechsler, David. Wechsler Intelligence Scale for Children: *Manual.* New York: Psychological Corp., 1949.

THE WECHSLER-BELLEVUE TEST AND THE DIAGNOSIS OF PSYCHOPATHIC PERSONALITY

MILTON S. GURVITZ

Hillside Hospital, Bellerose, N. Y.

Problem

The purpose of this study is to investigate the existence of a psychometric pattern for adult, white, male psychopaths resident in a penitentiary on the Wechsler-Bellevue test. The present study is limited to white males because the psychometric pattern of the female psychopath may differ and psychopaths of other races may introduce additional variables due to their unequal social status in our culture. The age of the population will be restricted to the age range of 19 to 35 since it is between these ages that the clinical picture of psychopathy reaches its sharpest focus and greatest incidence[6], and because the Wechsler-Bellevue utilizes that portion of the standardizing population that lies between the ages of 20 to 34 to weight the subtests. The intelligence range will be restricted to I. Q. 80 to 120 since above and below these limits the method of weighting the subtests and determining the I. Q. allows too great a possibility for both chance and systematic error.[11] Other studies exist which have set up patterns for psychopaths on the Wechsler-Bellevue, but clinical use has not confirmed these results in this writer's experience. The results of the only followup studies to date have been negative.[9, 10].

Method

The experimental population is composed of 84 white, male inmates of a Federal Penitentiary, ranging in age from 19 to 35, and with a Wechsler-Bellevue I. Q. in the range 80-119. On admission they were diagnosed as being psychopathic personalities and this diag-

nosis was sustained for at least one year. In addition to a clinical diagnosis, each experimental subject met the extremely severe criteria developed by Lindner to meet just such an experimental need.[5]

Control Group I consists of an equal number of inmates of the same penitentiary, diagnosed on their admission as: "Without mental disease" as the result of a detailed case history and clinical interview, and who manifested stable behavior during the entire period of their incarceration. Every member of this group was a first offender. As both the Experimental Group and Control Group I are prison groups, it was necessary to find another control group that would be representative of a larger population than a prison universe in order to evaluate the applicability of the results to the general population. Dr. Wechsler kindly made available the standardizing population of his test for this purpose. For purposes of conserving space, the Experimental Group will be referred to as: "X", Control Group I as: "C^1" and Control Group II as: "C^2".

The subjects were matched on an individual basis for the factors of race, sex, education, I. Q., age, and occupational status. All three of the above groups are similar in that they are white males. For each psychopath, a subject was chosen from each of the control groups that would closely match him for the remaining four factors of age, I. Q., education, and occupational status. It was not possible to match as rigorously with Control Group II as with Control Group I since the number of cases available to choose from was smaller. Tables 1 and 2 indicate the closeness with which orderly, paired matching was achieved. The high correlation coefficients are in reality underestimates due to the restricted range of the population.

Statistical Treatment Of Data. Means and sigmas for all quotient and subtest values were calculated. Statistical examination for differences between such groups as are used in this experiment depend on the normality of the distribution. All of the quotient and subtest distributions of the Experimental Group were tested for normality with the Chi square technique. Normally distributed groups were compared with the *t* tech-

TABLE 1
OCCUPATIONAL STATUS OF EXPERIMENTAL GROUP AND CONTROL GROUPS I & II

Occupational Group	X	C_1	C_2
Inexperienced or student	10	10	10
Professional	0	1	5
Farm owners and operators	0	0	0
Proprietors and managers	1	0	3
Clerical and sales	13	13	19
Craftsmen and foremen	1	1	4
Operatives	18	18	17
Domestic Service	0	0	0
Protective Service	15	15	0
Service Workers	3	3	3
Farm laborers and foremen	1	1	0
Laborers except farm and mine	22	22	23
N	84	84	84

TABLE 2
CORRELATION BETWEEN GROUPS ON VARIOUS MATCHING FACTORS

Criterion	X–C_1	X–C_2	C_1–C_2
Wechsler-Bellevue I. Q.	.95	.95	.92
Education	.95	.97	.97
Age	.86	.49	.46

Details of the matching procedure will be found in item 3 of the bibiliography.

nique in which the standard error of the difference between the means was computed directly from the intra-pair differences so that full account was taken of the correlation between the groups. Where the divergence from normality was significant, it appeared wiser to use the Chi square technique to compare the two groups.

Results

Testing of the quotient and subtest values for normality revealed that none of the quotient values were normally distributed, while three were doubtful. It then appeared that the wisest procedure was to use Chi square with all the values and to check those sufficiently normally distributed by the *t* technique.

Table 3

Significance Of Difference Between Means For Normally Distributed Wechsler-Bellevue Subtests

Subtest	*group*	D_M	*S.D.* d_M	*t*	*significance**
Information	X :C_1	.49	.27	1.82	nil
	X :C_2	.04	.30	.13	nil
	C_1 :C_2	.45	.28	1.61	nil
Arithmetic	X :C_1	.14	.40	.35	nil
	X :C_2	.70	.37	1.89	nil
	C_1 :C_2	.56	.38	1.47	nil
Block Design	X :C_1	.52	.39	1.33	nil
	X :C_2	.49	.34	1.44	nil
	C_1 :C_2	.03	.35	.09	nil
Digit Symbol	X :C_1	.15	.26	.58	nil
	X :C_2	.48	.28	1.71	nil
	C_1 :C_2	.63	.33	1.91	nil

* 83 degrees of freedom

Table 4

Significance of Wechsler-Bellevue Chi Square Distribution

Criterion	*Group*	*Degree of freedom*	*Chi Square*	*Significance*
Full Scale I. Q.	X :C_1	5	.23	.99
	X :C_2	5	.52	.99
	C_1 :C_2	5	.56	.98
Verbal I. Q.	X :C_1	5	2.52	.70–.80
	X :C_2	5	2.04	.80–.90
	C_1 :C_2	5	.52	.99
Performance I. Q.	X :C_1	5	2.74	.70–.80
	X :C_2	5	2.45	.70–.80
	C_1 :C_2	5	3.23	.50–.70
Information	X :C_1	3	.88	.90–.95
	X :C_2	3	.90	.85
	C_1 :C_2	3	2.04	.60
Comprehension	X :C_1	3	.64	.80–.90
	X :C_2	3	1.12	.70–.80
	C_1 :C_2	3	1.92	.60
Arithmetic	X :C_1	3	1.72	.60
	X C_2	3	2.98	.60
	C_1 C_2	3	4.62	.20
Digits	X :C_1	3	.50	.90
	X C_2	3	1.14	.70–.80
	C_1 :C_2	3	1.50	.60–.70
Similarities*	X :C_1	3	2.44	.40–.50

TABLE 4 (CONCLUDED)
SIGNIFICANCE OF WECHSLER-BELLEVUE CHI SQUARE DISTRIBUTION

Criterion	*Grouw*	*Degree of freedom*	*Chi Square*	*Significance*
Picture Completion	X :C_1	2	.04	.98
	X :C_2	2	.10	.95
	C_1 :C_2	2	.22	.80–.90
Picture Arrangement	X :C_1	2	1.18	.50–.70
	X :C_2	2	.32	.85
	C_1 :C_2	2	1.74	.30–.50
Object Assembly	X :C_1	2	.56	.70–.80
	X :C_2	2	3.08	.20–.30
	C_1 :C_2	2	1.22	.50–.70
Block Design	X :C_1	2	2.84	.20–.30
	X :C_2	2	1.26	.50–.70
	C_1 :C_2	2	.86	.50–.70
Digit Symbol	X :C_1	2	2.66	.20–.30
	X :C_2	2	2.08	.40
	C_1 :C_2	2	.08	.95–.98

* Only about a third of Wechsler's population had this subtest.

When tested by both the Chi Square and *t* techniques, none of the groups showed any differences of significance on either the quotient values of the Wechsler-Bellevue or its subtests. It must be concluded, therefore, that when the factors of sex, race, age, I. Q., education, and occupational status are held constant that:

1. There is no characteristic pattern for this group of psychopaths on the Wechsler-Bellevue Test.
2. The test patterns of the Wechsler-Bellevue for adult psychopaths and adult "normal" prisoners appears to be identical to a group of randomly selected subjects who were used to standardize the Wechsler-Bellevue Test.

The negative conclusions of this study regarding a psychopathic pattern on the Wechsler-Bellevue test are in agreement with the evaluative studies of Strother[9] and Van Vorst[10]. These conclusions are in sharp disagreement with the conclusions of Wechsler,[11] Levi[4] and Schafer.[8] As both Wechsler and Schafer derived their psychopathic patterns on the Wechsler-Bellevue from clinical experience with adolescents, it is not possible to subject the basis of their conclusions to scrutiny, but it would appear that the extension of these patterns to adults does not appear valid. This would apply equally as well to Levi's conclusions although he did not claim them to be applicable to adults.

SUMMARY

The purpose of this study was to discover whether there is a psychometric pattern for adult, white, male psychopathic personalities on the Wechsler-Bellevue test. The personnel consisted of three groups of 84 subjects: a psychopathic prison group; a control group of normal prisoners from the same institution; and a control group from the Wechsler-Bellevue standardizing population. The groups were equated for the factors of age, education, I. Q., economic status, race and sex. Quotient and subtest distributions were tested for normality of distribution, with only a minority proving to be normally distributed. Significant differences between groups on the various subtest and quotient values were calculated by Chi square for all of the values and *t* was also employed for those normally distributed. No characteristic pattern for psychopaths was found and the patterns proposed by Levi, Wechsler and Schafer were not found to be applicable to adults.

References

1. Bender, L. Psychopathic behavior disorders in children. In: *Handbook of Correctional Psychology.* New York: Philosophical Library, 1947.
2. Franklin, J. Discriminative value and patterns of the Wechsler-Bellevue Scale in the examination of delinquent Negro boys. *Educ. Psychol. Meas.*, 5, 1945, pp. 71-85.
3. Gurvitz, M. S. *The Experimental Application of Wechsler-Bellevue Type Tests In An Attempt To Discriminate And Diagnose Psychopathic Personality Types Resident In a Penal Institution.* Unpublished Doctoral Thesis, New York University, 1949.
4. Levi, J. A. *A Psychometric Pattern of The Adolescent Psychopath.* Unpublished Doctoral Thesis, New York University, 1943.
5. Linder, R. M. A Formulation of psychopathic personality. *Psychiatry*, 7, 1944, pp. 59-63.
6. Linder, R. M. The Rorschach test and the diagnosis of psychopathic personality. *J. Crim. Psychopathol.*, 5, 1943, pp. 69-93.
7. Machover, S. *Cultural and Racial Variations in Patterns of Intellect.* New York: Teachers College Contributions to Education No. 875, 1943.
8. Schafer, R. *The Clinical Application of Psychological Tests.* New York: International Universities Press, 1948, pp. 340-341.
9. Strother, C. The performance of psychopaths on the Wechsler-Bellevue test. *Proceedings of the Iowa Academy of Science*, 51, 1944, pp. 397-400.
10. Van Vorst, R. "Abstract." *Psychol. Bull.*, 41, 1943, pp. 583.
11. Wechsler, D. *The Measurement of Adult Intelligence.* Baltimore: Williams and Wilkins, 1944, Third Edition.

A STUDY OF THE DIFFERENTIAL RESPONSES ON THE VOCABULARY SUB-TEST OF THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN

CHALMERS L. STACEY AND BERNARD PORTNOY

Syracuse University

Problem

Gerstein[1] has suggested the value of placing more emphasis on the relationship between the use of language and the underlying cognitive process. That is, instead of merely determining whether a definition of a word given on a vocabulary test is correct or incorrect, one should, in addition, endeavor to differentiate the type of definition credited as correct. Differential patterns of response to intelligence tests have been investigated extensively in the hope that these patterns might aid in the diagnosis of functional and mental disturbances. Gerstein suggests that these differential patterns might emerge in responses to items on a vocabulary test, e.g. the vocabulary section of the Wechsler Intelligence Scale for Children.

Reichard, Schneider, and Rapaport[3], using the Goldstein-Weigl[2] Color Sorting Test as an instrument to study concept formation in children, concluded that three methods of forming concepts exist: (a) the concretistic, (b) the functional, and (c) the conceptual. Gerstein assumed that the same three methods used in solving a performance problem may be used in classifying definitional levels produced by verbal stimuli, and suggested the following classification: (a) the concretistic or descriptive method, which is considered the most primitive level of verbalized definitions, e.g. "An apple is red.", (b) the functional or usage method which is a more complex method of verbalization, e.g. "A bicycle is something you ride.", (c) the categorical or conceptual method which represents an abstract method of definition. This type of definition is considered more complex than the preceeding two, e.g. "A donkey is an animal." It is the aim of this paper to test the above hypothesis.

Method

Two groups of children were used as subjects. One group was classified as mental defective and the other borderline, according to the Wechsler Intelligence Scale for Children. Their characteristics can be seen in table I. The responses of each child to the words in the vocabulary section of the Wechsler Intelligence Scale for Children were re-

TABLE I. PERFORMANCE OF BORDERLINE AND MENTAL DEFECTIVE SUBJECTS ON W. I. S. C.

	N	(in years) Chronological Age		Verbal Quotient		Performance Quotient		Total I.Q.	
		Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
MENTAL DEFECTIVE (I.Q. 50-65)	24	10.77	2.53	60.96	4.60	63.08	7.50	58.42	4.74
BORDERLINE (I.Q. 66-79)	27	12.12	2.01	69.70	5.39	76.52	6.74	70.44	3.55

corded verbatim. Each response was entered on a chart under the appropriate category, the descriptive, the functional, and the categorical. Only nouns were used. The total number of words in each category was computed. The proportion of each type of response (descriptive, functional, and categorical) to the total number of responses was then calculated for each of the two groups. Critical ratios were computed between the three pairs of proportions. The results are presented in table II.

TABLE II. FREQUENCY AND PROPORTION OF LEVELS OF RESPONSE TO VOCABULARY SECTION OF THE W. I. S. C.

Category	Mental Defective		Borderline		Diff in Prop.	C. R.
	Frequency	Proportions	Frequency	Proportions		
Descriptive	10	.054	38	.144	.09	3.00
Functional	166	.897	212	.802	.094	2.69
Categorical	9	.049	14	.053	.004	.19
TOTAL	185		264			

RESULTS

1. The proportions of descriptive responses made by the two groups were significantly different at the one percent level. The Borderline group gave a significantly greater number of descriptive responses.
2. The proportion of functional responses in the two groups were significantly different at the one percent level. There were more functional responses in the mental defective group.
3. No significant difference was found for the categorical responses ($p = .85$) although the tendency for the Borderline group to give a greater number of categorical responses was present.

SUMMARY

The responses of two groups of children, one mentally defective and one Borderline, were analysed in terms of total I.Q. on the Wechsler Intelligence Scale for Children, and as to the proportion of their vocabulary response that fell into the descriptive, functional, or categorical categories. Contrary to what was expected, a significantly greater proportion of descriptive responses were found to occur in the Borderline group than in the mental defective, and a significantly greater proportion of functional responses were found to occur in the mental defective group than in the borderline group. Whether this result is restricted to our sample remains to be seen. However, the results do leave in

doubt the fact that the functional response is of a higher level than the descriptive. Further study is necessary to make these findings conclusive.

References

1. Gerstein, Reva A. A suggested method for analysing and extending the use of Bellevue-Wechsler Vocabulary responses. *J. Consul. Psychol.*, 1949, 13, 366-370.
2. Goldstein, K. and Scheerer, M. Abstract and concrete behavior: an experimental study with special tests. *Psychol. Monogr.*, 1941, No. 2.
3. Reichard, S., Schneider, M., and Rapaport, D. The development of concept formation in children. *Amer. J. Orthopsychiat.*, 1944, 156-161.

REPORT ON A METHOD OF PREDICTING SUCCESS IN PSYCHOTHERAPY

theodore H. Blau

The Pennsylvania State College

Introduction

Only recently has widespread interest in the objective evaluation of therapeutic processes arisen. A process describing the nature of nondirective psychotherapy was reported in 1945 by Snyder[5]. The course of therapy was described in terms of the verbalizations of both the client and the counselor, using a system of response categories in the description. In 1947, Dollard and Mowrer[1] described a method by which client statements made during the course of therapy are classified for the calculation of a "discomfort-relief quotient" (DRQ) which seemed useful in describing therapy progress. Raimy[3], in 1948, reported the development of a "positive-negative-ambivalent quotient" (PNAvQ), accomplishing a purpose similar to that of the DRQ. In a later publication, Kauffman and Raimy[2] reported research indicating that the "DRQ" and "PNAvQ" were highly similar in nature and purpose.

With the establishment of high reliability in response classification by Snyder[5] and later by Seeman[4] and others, the possibility that statements made by the client during the first interview of therapy might be indicative of success in therapy was suggested. In order to investigate the value of the first interview for the prediction of therapeutic success, a preliminary research was undertaken.

Procedure

Ten of Snyder's[5] categories were considered as the best available method for the classification of client statements in therapy interviews. Table 1 lists the response categories and their abbreviations.

Table 1. Client Response Categories and Their Abbreviations

Response Category	Abbreviation
Positive Attitude Toward the Self	PAS
Ambivalent Attitude Toward the Self	AmS
Negative Attitude Toward the Self	NAS
Positive Attitude Toward the Counselor	PAC
Ambivalent Attitude Toward the Counselor	AmC
Negative Attitude Toward the Counselor	NAC
Positive Attitude Toward Others	PAO
Ambivalent Attitude Toward Others	AmO
Negative Attitude Toward Others	NAO
Miscellaneous	M

Seven completed cases of personality adjustment problems from the files of the Psychological Clinic of the Pennsylvania State College were selected as the experimental sample for this study. All cases had been treated according to non-directive procedures. The number of interviews in each case ranged from two to eighteen, with an average of six interviews per case. The verbatim typed transcriptions of the first interview of each case were obtained from electrical recordings of these interviews. The client statements were classified in terms of the categories listed in Table 1. Each client statement was assigned a "valence" and a "direction" according to the following criteria (abbreviated for this report):

Valence. Valence is defined as the value or attitude of acceptability placed on anything by the client in his verbalizations. The following arbitrary valences are considered in this study:

1. *Positive.* A positive valence is defined as a value of "good" or "acceptable" expressed by the client in connection with any directed statement.
2. *Ambivalent.* An ambivalent valence is defined as the co-existance of a positive and a negative valence in connection with any directed statement. Client statements of conflict situations are classified under this category.
3. *Negative.* A negative valence is defined as a value of "bad" or "non-acceptable" expressed by the client in connection with any directed statement. It is the opposite of the positive valence.

Direction. Direction is defined as the object toward which the client expresses his valenced statements. The following directions are considered in this study:

1. *Self.* A client statement is self-directed when the client verbalizes anything in which he is the recipient of a valence. Verbalization of the problem is to be considered a self-reference unless specifically directed toward the counselor or others.
2. *Counselor.* A client statement is counselor-directed whenever the statement is valenced, and primarily concerns the counselor, counseling, or the counseling situation.
3. *Others.* A client statement is other-directed whenever the statement is valenced, and primarily concerns anyone or anything other than the self or the counselor.

Miscellaneous. This category is reserved for client statements which evidence no valence, for example:

> Client: Today is January the twenty-third.

A client statement, or response, is defined as all words spoken by the client between any two consecutive counselor statements. The counselor statement immediately preceding any client statement must be considered in the classification of that client statement.

A partial check on the reliability of the experimenter's classifications was taken in this preliminary study. Ten days after the original classifications were made, the client statements (a total of 407) in the seven first interviews were re-classified, with no reference to the original classifications. It was found that the two groups of classifications were 95% in agreement. This figure is probably weighted positively because of practice effect, short time interval between rating sessions, and the single-judge factor.

Each case was rated for "success" by the original counselor after re-reading his own case summary. A rating scale of "therapy success", as developed by an organized research group made up of eight candidates for the Ph.D. in clinical psychology, was used in the evaluation of each case. As a result of these ratings, each case was assigned a "success value". In order to estimate the reliability of this rating technique, the counselors re-rated their cases approximately four weeks after the original evaluation, using identical procedures. Using an ungrouped data formula, the first and second ratings were compared, the re-

sulting coefficient of correlation being .94, which is significant below the 0.01 level of confidence. This reliability coefficient must be viewed with the same caution applicable to the above classification data.

After considerable review of the literature, a theoretical basis for the prediction of success in psychotherapy was formulated as follows:

1. It is proposed that the self categories of client statements in the first interview of psychotherapy are predictors of therapy success.

2. It is further proposed that valenced self statements are measures of behavior acceptabilty habits, positive valences indicating accepted behavior habits, negative valences indicating non-accepted behavior habits, and ambivalent valences indicating behavior habits in the process of valence conflict or change.

3. On the basis of the "self concept" as developed by Raimy[3] and the preceding premises, it would be expected that the greater the number of positively-valenced (PAS) and ambivalently-valenced (AmS), and the fewer the number of negatively-valenced (NAS) self statements present in the first interview of psychotherapy, the higher will be the prediction for therapy success.

In order to test the preceding assumptions on the data in this study, the following formula was derived:

$$PI = PAS\% + AmS\% - NAS\%$$

where PI stands for "predictive index".

Results

The single classification categories, as listed in Table 1, were tabulated for each first interview in the cases sampled. In each case, the tabulated classification categories were converted to percentage of total client statements for the particular interview in order to equate the interviews for the number of client statements per interview.

Applying the above formula, a predictive index was obtained from the first interview of each case. These indices, as well as success values for each case, may be seen in Table 2.

The predictive indices were correlated against success values using an ungrouped data formula for r. The resulting coefficient of correlation was found to be .90, which is significant below the 0.01 level of confidence.

Summary

1. The classification of client statements in the first interviews of seven personality adjustment cases was investigated as a method of predicting the success of nondirective psychotherapy.

2. On the basis of previous research in the process of psychotherapy, a formula for the prediction of success in therapy was derived. This formula was tested against the cases sampled with a resulting coefficient of correlation of .90 between the prediction formula values and ratings of therapy success.

3. Since this investigation involved a very limited experimental sample, the results must be considered as merely indicative of the value of the prediction

Table 2. Predictive Indices Obtained from First Interviews and Success Values for Each

Case number	Predictive index	Success value
1	11	34
2	10	31
3	0	23
4	−6	18
5	−21	15
6	4	21
7	−10	17

formula. A research project involving a larger case sample and more adequate controls is now being planned in order to further test the hypotheses developed in the present investigation.

REFERENCES

1. DOLLARD, J. and MOWRER, O. H. A method for measuring tension in written documents. *J. abn. and soc. Psychol.*, 197, *42*, 3-32.
2. KAUFFMAN, P. E. and RAIMY, V. C. Two methods of assessing therapeutic progress. *J. abn. and soc. Psychol.*, 1949, *44*, 379-385.
3. RAIMY, V. C. Self-reference in counseling interviews. *J. consult. Psychol.*, 1948, *12*, 153-163.
4. SEEMAN, J. A study of the process of nondirective therapy. *J. consult. Psychol.*, 1949, *13*, 157-168.
5. SNYDER, W. U. An investigation in the nature of nondirective psychotherapy. *J. gen. Psychol.*, 1945, *33*, 193-223.

DIFFERENTIAL SCORING OF THE ROSENZWEIG PICTURE-FRUSTRATION STUDY

GEORGE SPACHE

Chappaqua, N. Y.

Use of the children's edition of the Rosenzweig Picture - Frustration Study has led to the formulation of a technic to permit the examiner to differentiate between the child's reaction to frustration in child-child and adult-child relationships. Described briefly, the P-F study is a series of 24 cartoon-like pictures depicting conflict between children or between children and adults. Response is secured by asking the child to supply the probable remarks of one of the figures in each conflict situation[3].

As scored according to Rosenzweig's instructions[2], the direction as well as the type of aggressiveness of the respondent is evaluated. Three scores are distinguished under direction of aggressiveness: extrapunitiveness (E), in which aggression is directed outwardly onto the environment; intropunitiveness (I), in which it is turned back upon self, as in accepting blame; and impunitiveness (M), in which aggressive action is avoided by minimizing or ignoring the frustration. The type of reaction also involves three scores: obstacle dominance (O-D), in which the incident causing the frustration is emphasized; ego defense (E-D), in which the personalities of the individuals involved in the conflict are stressed; and need-persistance (N-P), in which the solution of the problem is given paramount importance. In addition, a group conformity rating (GCR) is obtained in terms of the percent of agreement of the individual with the modal responses to each item as found in Rosenzweig's standardization data.

In the present Children's Form of the P-F study, 13 of the pictures represent conflict between an adult and a child. The remaining eleven involve conflict between children. It is true that a general concept of the child's responses to frustration and conflict may be obtained from his scores on the entire 24, but our use of the test in study of the maladjustments of elementary-school children soon showed that this generalized scoring technic, failed to differentiate among children whose tendencies varied according to the individual with whom they were in conflict. Many seriously maladjusted children differ considerably in the nature of their responses to the conflict situations with adults from their responses to those incidents involving other children.

The description of a particular case may serve to illustrate this point.

B. C. is a nine-year-old, fourth grader of average intelligence. His academic achievement is poor: his reading, quite retarded despite remedial training. His teacher complains of his lack of interest and effort; his inability to form associations with classmates and his withdrawal from all group activities. His play behavior is to imagine himself a car and to imitate the steering, shifting, backing, etc., constantly. This solitary play goes on day after day and enables him to avoid close contact with the remainder of the class.

In preliminary diagnosis, this would appear to be a problem of social adjustment, of withdrawal from unsuccessful reality into a fantasy world that supplies the needed feelings of success and acceptance. In general, the analysis of responses to the entire P-F supports this impression. His results were: E 56%; I 8%; M 36%; O-D 10%; E-D 69%; N-P 21%; GCR 69%. The only significant trend in his responses was N-P —.60. We would interpret these data in this fashion:

On the whole, his responses are quite similar to those of children of his age, *i.e.*, his aggressive tendencies are expressed in a reasonably conventionalized manner (GCR). While the outward expression of his aggressiveness (E) is high, it is within the normal range for a child of his age and he resembles such children in the more obvious types of aggressive behavior. Insight into his own contribution to conflict, and the ability to accept blame or censure (I) are markedly weak. As a result, he is sensitive to criticism, defensive and resentful of interference by others (E-D) to an abnormal degree. His responses to frustration are largely direct and personal. He seldom spontaneously accepts blame or offers to make amends. When directly accused or blamed (as in Rosenzweig's "super ego-blocking" situations), he tends to come back with vigorous denial or counter aggressiveness. As the length of his association with others increases, there is a tendency for him to be less defensive and to seek mutually acceptable solutions for conflict. (N-P trend).

When B. C.'s responses to conflict with children or adults are differentiated, a clearer and more accurate picture of his difficulties is obtained. These responses are summarized on the Supplementary Record Blank we have devised for this purpose (Fig. 1).

Using this approach, the indications, in our opinion, are as follows:

In dealing with adults, B. C. is quite different than most children of his age. He is markedly more aggressive, resistant and defensive. He is critical of adult interference and sensitive to personal affront. He has little understanding of his own part in these conflicts and responds to adult censure with denial and resentment. Thus he may fail to respond to attempts to motivate him or to function at the level of which he is capable, academically speaking. On the whole, his responses differ seriously from the conventional pattern of the average child in his dealings with adults.

In dealing with other children, B. C.'s attitudes are strongly repressed and conventionalized. He is significantly less outspokenly aggressive than the average child of his age. He exercises considerable self-restraint, avoids conflict by peace-making attempts and by glossing over or ignoring the aggressions of others. Thus his adjustment is largely an adaptive one involving little or no active effort on his part. There are, however, favorable trends toward greater aggressiveness or self-assertiveness, and toward more active seeking of solutions to situations of conflict.

It is apparent from these analyses that the two approaches to the scoring of B. C.'s. responses yield several significantly different interpretations. We find, for example, that the degree of conventionalization of responses is excessive and adjustive with children, but much less with adults, rather than being similar to the average child as indicated in the entire P-F. His outward and obvious expressions of aggressiveness are directed largely toward adults, as are his defensiveness, sensitivity and resistance. Toward children he is adjustive, severely repressed, and his social role is largely a negative one. These are significant and accurate insights which we would not have gained without the detailed analysis of his responses. Rather than appearing as a reasonably well-adjusted child except for his lack of insight and defensiveness, B. C. is now seen as badly adjusted to both children and adults with opposite tendencies in each type of situation. The detailed analysis also gives an insight into his lack of academic progress and into his solitary play habits.

We realize that the method of comparing the child's responses with the modal answers is crude and that there are no normative data available for comparing the GCR's in the two situations. But despite these handicaps, the Supplementary Record Blank is an aid to more careful analysis and understanding of the child's responses to frustration and conflict. As we have tried to show in the case study, children present different facets of their personalities in their dealings with their peers and with adults. The original method of scoring the test as an entire unit tends to conceal these differences. We have found the more detailed analysis is more accurate and insightful with those children who differ in their adjustment in these distinct milieus.

SUPPLEMENTARY RECORD BLANK—CHILDREN'S FORM ROSENZWEIG P–F STUDY

GEORGE SPACHE, One Salem Road, Chappaqua, N. Y.

For analysis of situations involving adult-child and child-child conflicts. Enter scores on items of adult-child conflict for which modal responses are given below. Summarize these and the eleven items depicting child-child conflict under profile.

Name **B. C.** Date Age **Nine**

	O-D	*E-D*	*N-P*
1.	E'	**E**E	e
4.		**M**	e
5.	**E'** E'		e
7.		**E** I	
10.		**M** I	
11.			**m** m
13.		**E** E,I	
14.	M'	**E**	
17.			**m** m
19.		**E** I	
22.		**E** I	
23.	M'	M	**e** e
24.			**m** m

Profile

	O-D	*E-D*	*N-P*	*A-C*	*C-C*
E	**1** (2) **0** (1)	**6** (2) **4½**6(1)	**1** 1(3) **1**	2–5 **8**	**5½** 6–8
I	**0** **1** (1)	**0** 4(1) **1** (1)	**0** **0** (1)	4–5 **0**	2 0–3
M	**0** 1(1) **½**	2 (1) **3** 2(1)	**3** 3 **0**	4–5 **5**	**3½** 2–3
A–C	**1** 1–4	**8** 5–7	**4** 4–7		
C–C	**1½** 0–2	**8½** 8–11	**1** 0–1		

The A–C and C–C columns indicate the range of modal or expected answers in Adult-Child and Child-Child situations. They are the sums of the modal responses and the alternate answers in the parentheses.

Trends

	A–C	*C–C*
1.	**None**	**None**
2.	" (M–.33) ←	**E—.33** ←
3.	"	**None** (e+. 50) ←
4.	" (M– .50) ←	" (M–.50–.60) ←
5.	"	**N.P. – .50** ←

7/13 = **54**% GCR in Adult–Child
9.5/11 = **86**% GCR in Child–Child

BIBLIOGRAPHY

1. ROSENZWEIG, SAUL. *The Children's Form of the Picture-Frustration Study.* 4562 Scott Avenue, St. Louis, Mo.: The author.
2. ROSENZWEIG, SAUL. *The Children's Form of the Picture-Frustration Study: Criteria for Interpretation.* St. Louis, Mo.: The author.
3. ROSENZWEIG, SAUL. The Picture Association Method and Its Application in a Study of Reactions to Frustration. *J. Personality,* 1945, 14, 3-23.

A COMPARISON OF TWO METHODS OF ADMINISTERING THE DIGIT SPAN TEST

RICHARD L. NEWTON

University of Pittsburgh

Problem

For many years the digit span test has been used as one of the items included in intelligence tests, namely as a measure of the memory span. The two leading individual intelligence tests in general use at the present time, the Revised Stanford Binet Tests of Intelligence[3] and the Wechsler-Bellevue Adult Intelligence Scale[4], use the digit span test for this purpose.

Concerning the administration of the digit span test, Terman and Merrill [3], *p. 79* state, "Pronounce the digits distinctly and with perfectly uniform emphasis at the rate of one per second." Wechsler[4], *p. 176* gives these instructions, "Read digits at the rate of one per second." It will be noted that Wechsler makes no stipulation concerning the emphasis to be given in administering the series. Neither manual of the two scales makes any statement concerning emphasis in the administration process used in standardizing the tests.

It was discovered that in clinical use there are two major methods of administration in reference to the emphasis given the last digit of each series in a digit span test. Some examiners feel that the pitch of the voice should be held constant throughout the series of digits. Other examiners feel that the pitch of the voice should be lowered on the last digit of a series thus signifying to the subject that he may begin immediately to repeat the series. The purpose of this study is to conduct a controlled experiment to ascertain whether the method of administration alters significantly the results obtained on digit span tests.

Method

The Wechsler-Bellevue Adult Intelligence Scale Form I was administered to twenty-eight patients at the Veterans Administration Hospital in Aspinwall, Pennsylvania. All of the patients used in the study were referred to the psychological service for diagnostic testing. It is assumed that these twenty-eight patients are a fairly representative and random sample of the psychiatric patients at the hospital in that they were selected consecutively as they were referred for psychological testing. It must be noted that severely disturbed psychotic patients are not usually subjected to psychological testing and none were included in the present sample.

The Wechsler-Bellevue Adult Intelligence Scale, Form I[4] was administered to each patient with the order of the subtests being held constant for each patient. Also each patient was administered an additional digit span test, namely, those digit series used in the Babcock-Levy Test[1]. This was done so that one set of digits could be administered with the pitch of the voice being lowered on the last digit of each series, and the other set of digits administered with the pitch of the voice being held constant on the last digit of each series. Thus results were obtained for each patient for each method of administration.

Table I shows the order of presentation to four subjects. This order of presentation was repeated for the remaining twenty subjects.

In following the procedure described above, certain other variables were controlled as closely as possible. Each sub-

Table I

Sample of Counter Balanced Procedure Used

Subject	*1st Digit Test*	*Pitch of Voice*	*2nd Digit Test*	*Pitch of Voice*
1	W–B*	Lowered	B–L	Constant
2	W–B	Constant	B–L	Lowered
3	B–L**	Lowered	W–B	Constant
4	B–L	Constant	W–B	Lowered

* W–B refers to the Wechsler-Bellevue digit span test

** B–L refers to the Babcock-Levy Digit span test

ject was given the same instructions concerning the test. All subjects were tested under the same physical conditions and in the same room. The same sub-tests were administered between the two digit span tests for each subject. All of the digits were presented at the same rate for each patient, and the examiner attempted to eliminate any uncontrolled variation in his method of presentation among the patients. Each patient was ascertained to have normal auditory acuity and no patient had been administered a Wechsler-Bellevue Test for at least one year previous to the present testing. The age range for the entire group was from 19 to 59. The intelligence ranged from a Wechsler Bellevue I. Q. of 74 to 124. It was found that these ranges were fairly evenly distributed over the four divisions required by the research procedure.

A "t" value between the two methods of administration for the digit span test as a whole (both forward and reversed) was computed according to the method described by Peters and Van Voorhis[2]. A further breakdown involved computing "t" values by the same method for the digits forward and the digits reversed. Another "t" value was computed to determine the significance of any difference in difficulty between the two different sets of digits used. Each of these "t" values was examined for significance at the one per cent level of confidence taking into account the appropriate degrees of freedom.

Results

The "t" value (3.58)* obtained for the digit span test as a whole was significant far beyond the one per cent level and in the direction indicating that lowering the pitch of voice on the last digit of each series enables the subject to recall more digits correctly.

The "t" value (3.90)* obtained for the digits forward was significant far beyond the one per cent level and in the direction that indicates that lowering the pitch of the voice in the last digit of each series enables the subject to obtain a much better score on digits forward.

The "t" value (1.45)* obtained for the digits reversed was not significant. This finding would suggest that the method of administration does not affect the results of the digits reversed in any appreciable manner.

The possible effect of a difference in difficulty between the two sets of digits was cancelled out by the use of the counter balanced procedure used in the administration. However, it was felt that information concerning the difficulty of the two sets of digits would be valuable and interesting. For that reason a "t" value (1.21) was computed for a difference in difficulty between the Wechsler-Bellevue digits and the Babcock-Levy digits, and was found to be not significant at either the one or five per cent level of confidence.

Table II contains the mean differences, standard deviations, standard errors, and "t" values which show the differences resulting from the two methods of administration.

Discussion

These results indicate that the method of administration of digit span tests significantly affects the results on scores obtained. Lowering the pitch of the voice on the last digit of each series enables the subject to recall more digits than if the pitch of the voice is held constant. These results pertain only to the digits forward, the method of administration not significantly affecting the results for the digits reversed. Con-

*A "t" value of 2.77 is required for significance at the one per cent level.

Table II

Comparisons of Different Methods of Administration

	Test as Whole	*Digits Forward*	*Digits Reversed*
Mean Difference between methods	1.11	.82	.29
Sigma	1.59	1.09	1.03
S. E. of Sigma	.31	.21	.20
"t" value	3.58*	3.90*	1.45

cerning the recall of digits forward it is felt that the lower pitch of the voice on the last digit is a signal to the subject that he may begin immediately to repeat the series and thus no time interval ensues before recall. When the pitch of the voice is not lowered on the last digit of a series, the subject waits for the next digit to be given, thus, time elapses before active recall begins and this possibly interferes with the accuracy of recall.

It has been ascertained that subjects normally do not recall as many digits reversed as are recalled forward, consequently, fewer series of digits are necessary to ascertain the maximum level of accurate recall for digits reversed. It may be that these shorter series (relative to each subject's performance on digits forward) are not affected by the time lapse between the termination of administration and the beginning of response. In addition, most subjects delay their response for a short period in order to memorize the series forward before attempting to repeat the numbers in reversed order. This usual delay may rule out the effect of the time lapse mentioned above and thus also contribute to the negative results for digits reversed.

It is pointed out that such differences in raw scores on the digit span test affect the weighted scores of the subjects and, in turn, affect the intelligence quotients of the subjects in varying degrees depending upon their age and intelligence level.

The results of this study can best be illustrated by their application to the intelligence quotient values contained in the tables cited by Wechsler[4] and Terman and Merrill[3]. Table 47 in Wechsler[4], *p. 228* is a table of weighted score values derived from the raw score values obtained on each of the sub-tests. Examination of this table reveals that a difference of one point in raw score for the digit span test produces a difference of at least one and frequently two points in the weighted score. Since the present study reports a mean difference of 1.11 points in raw score due to different methods of administration of the digit span test, it is evident that such a difference can affect a subject's weighted score by atl least two points. Wechsler's[4], *pp. 239-246* table of verbal intelligence quotients shows that two points in weighted score can affect the intelligence quotient by four points depending upon the age of the subject and his total weighted score. The minimum effect of two weighted score points is two intelligence quotient points, again depending upon the age and intellectual level of the subject.

In the Revised Stanford Binet Test of Intelligence, a difference in recall of one digit can result in either success or failure of that test item at any year level. In accord with the scoring procedures of the Revised Stanford Binet, each test counts as one month of mental age credit from years two to five, two months credit from years six to fourteen, two months credit at the Average Adult Level, four months credit at the Superior Adult Level I, five months credit at Superior Adult Level II, and six months credit at Superior Adult Level III. From the Intelligence Quotient Tables cited by Terman and Merrill[3], *p. 414ff* it can be shown that these mental age values can affect the intelligence quotient by as much as four points in each of the levels mentioned above.

This variation of intelligence quotient values resulting from different methods of administration of digit span tests points up the need of a consistent method of administration in accord with the standardization process which was used for every measuring instrument making use of the digit span test.

Summary

A comparison of two methods of administration of the digit span test of the Wechsler-Bellevue Adult Intelligence Scale, Form I was made using twenty-eight hospitalized psychiatric patients at the Veterans Administration Hospital in Aspinwall, Pa. The following results were obtained:

1. The pitch of the voice on the last digit of a series significantly affected the results obtained in the subjects' score. If the pitch of the voice was lowered the subjects recalled more digits accurately.
2. The pitch of the voice significantly affected the digits forward but did not have a significant effect upon the digits reversed.

3. There was not a significant difference in difficulty between the set of digits used by Wechsler and the set of digits used by Babcock and Levy.

BIBLIOGRAPHY

1. BABCOCK, H. and LEVY, L. *The Revised Examination for the Measurement of the Efficiency of Mental Functioning.* Chicago: C. H. Stoelting, 1942.
2. PETERS, C. C. and VAN VOORHIS, W. R. *Statistical Procedures and Their Mathematical Bases.* New York: McGraw-Hill, 1940.
3. TERMAN, L. M. and MERRILL, M. A. *Measuring Intelligence.* Boston: Houghton Mifflin, 1937.
4. WECHSLER, D. *The Measurement of Adult Intelligence.* Baltimore: Williams and Wilkins, 1944.

A NOTE REGARDING THE SUGGESTED USE OF THE BENDER VISUAL MOTOR GESTALT TEST AS A MEASURE OF SCHOOL READINESS

MARCELLA VIG BALDWIN

The Bureau for Colored Children
Philadelphia, Pennsylvania

INTRODUCTION

In a recent article in this journal[3], the Harrimans report their investigation of the thesis "that a child who is making satisfactory progress in reading must be at a maturational level whose sensory-perceptual-motor activities more closely resemble those of adults than they do the responses of pre-readers": more specifically, that successful young readers are able to reproduce the Bender Visual-Motor Gestalt Test figures with an accuracy more nearly approaching that of adults than children who have not yet learned to read. In essence, the Harrimans' investigation was a miniature repetition—miniature in terms both of age range studied and number of cases collected—of Bender's own attempt[1] to establish norms for these figures so that they might serve as a performance test of maturational level for children between the ages of 4 and 11. The Harrimans' data consisted of the Visual-Motor Gestalt Test productions of two groups of children, the first composed of nursery-school children whom they called "pre-readers", and the other of second-graders known to be making satisfactory progress in reading. Their findings were consistent with their expectations, and agree also with Bender's norms which portray so vividly the marked changes in the visual-motor gestalt function occurring between 4 and 7.

Bender has concluded from her studies of sensory aphasia that the visual-motor gestalt function is closely associated with language ability and suggested therefore that observed maturational changes in it might be of interest with regard to the development of language itself, particularly since her researches have shown that the age period when children conventionally are taught to read and write coincides with that in which the most rapid development of the gestalt function is taking place. The Harrimans have carried her suggestion farther, concluding from their data that the observed differences between the gestalt productions of the nursery school children and the second graders bear a fairly direct relationship to their reading skill or lack of it, even stating that perhaps the test may be used as a measure of the pupil's readiness to benefit from a "normal" curriculum.

While maturational level of the visual motor gestalt function as indicated by performance on the Bender may turn out to be *one* index of reading readiness, a causal relationship cannot be assumed on the basis of the Harriman study. The chief novelty of their approach was their emphasis on the fact that the younger children were not yet reading and that that the older ones were doing so. These authors, however, neglected to take account of the other momentous changes in general intellectual and physical de-

velopment taking place between the preschool years and second grade. Actually, on the basis of their data as presented there is no ground for assuming any closer relationship between the visual-motor gestalt function as revealed through the Bender and readiness for reading than that they are both manifestations of the larger developmental process. Their conclusion might, with equal logic, have been that because children whose *heights* more nearly approximate those of adults than of nursery school children are found to read, height may be taken as an index of the likelihood that a child will profit by the "normal" curriculum! Some weight might have been attached to their conclusions had an actual attempt to teach the younger group to read been unsuccessful; a crucial study, however, must involve a comparison of the Bender productions of readers and non-readers of similar age and school exposure.[1]

The following clinical observations are reported to re-emphasize that the prediction of reading aptitude is a complex matter perhaps never to be achieved with precision by any one instrument, and also to record a rather dramatic "miss" by the Bender. This last, insignificant as it seems taken alone, is the type of occurrence with which any adequate theory must be prepared to cope.

Case Report

Nancy and Pamela, two Negro girls, sisters, aged 12 and 13 respectively, who were wards of a child-placing agency, were referred for psychological study because of their failure to make an adequate adjustment in one of a series of foster homes. Their histories and the test results obtained will not be discussed in detail because the point to be made is a simple one. Suffice it to say that the complaints made by the foster parents were of a type frequently heard with regard to this group of foster children as they approach adolescence: insolence, "stealing" of food and small change, and general unreliability.

[1]Such a study is now in progress at the Bureau for Colored Children.

On the 1937 Stanford-Binet, Pamela obtained an I.Q. of 75 (M.A. 10-0) and Nancy an I.Q. of 58 (M.A. 7-4). Equally striking differences appeared between their Bender Visual Motor Gestalt Test performances, which were in the direction one would anticipate. Pamela, who found the task an enjoyable one, drew neat, logically arranged figures, all of which, with one questionable exception, would rate at the seven year level or above on Bender's norms.[2] Nancy, although she worked persistently, grumbled over the difficulty of the task and recognized the inferior quality of her productions but was unable to alter them for the better. Only two of her drawings could, even by a generous interpretation, be rated at the seven year level and three were no better than those of a five-year old, with the rest scoring at the six year level. Her drawings were arranged in a very irregular, crowded order. When they are analyzed in terms of Hutt's[4] categories, they show condensation and simplification, difficulty with closure effects and with overlapping figures, and failures in angulation. None of these characteristics appeared clearly in Pamela's figures with the exception of one case of failure in angulation, something that occurred four times among Nancy's drawings.[3]

Were one to predict the level of reading skill likely to be achieved by these two girls on the basis of the above test results, his estimate probably would be that Pamela would have difficulty in going beyond the third grade level or thereabouts because of her general intellectual limitations. Although now at adolescence she performs adequately on

[2]Bender has noted that there is relatively little change in the gestalt figures between the ages over seven, the chief addition being an improvement in obliquity and an increase in the number of combinations, with the result that the rating of productions above this age level becomes a much more difficult and more subjective matter than at the earlier ages.

[3]Emotional factors and their possible influence on the Bender results are being disregarded here since the case of Pamela may be considered as a "control"; objectively considered, both girls had been subjected to similar traumata and both were considered equally maladjusted by the foster parents and the case worker.

the Bender, it would seem unlikely that this alone could compensate for her retardation in other respects. Nancy's case particularly in the light of Harrimans' hypothesis, would seem to be even more clear-cut. While it is not uncommon for mentally retarded girls of this level to handle the mechanics of reading with some facility, her I.Q., coupled with her highly immature productions on the Bender, would make it seem unlikely that she could get much beyond first grade material.

Gray's Oral Reading Paragraphs[2] were used to determine the reading levels actually reached by the two girls. Pamela, who had handled the Bender with such ease, was found to be reading at the 3.2 Grade Level, which is not inconsistent with the above discussion. Nancy, on the other hand, whose Bender reproductions were infantile even considering her mental age, read at the 3.9 Grade Level! They both made similar errors of substitution, repetition, etc., on the early paragraphs with faulty word analysis not appearing until Paragraph 5. Pamela, as one would expect, comprehended the material better than Nancy; she could accurately report on what she had read in all the first four paragraphs while Nancy began to become confused even on the second. It will be noted that in this paper "reading" is being defined merely in terms of translating written symbols into spoken language, with little regard for the degree of understanding the translation once made!

Discussion

The cases reported above may be considered inappropriate for discussion in the present context because of their age, intellectual limitations and behavioral maladjustment. In a certain broad sense it is true that the Bender might have been used in a general way to evaluate school readiness here: considering her I.Q., it seems that although now at adolescence Pamela's Bender is of the adult type, at the time she entered school it may well have been in keeping with her mental rather than her chronological age level (as noted in an earlier footnote the Bender becomes less effective as a differentiating device at the later ages). Nancy with her severe retardation is not even now getting much out of her school exposure so the Bender again might be said to be consistent with the observed facts. Nevertheless, with regard to the specific point of reading aptitude, the Bender *has* "missed" here. It is clear that an intellectually retarded girl with the Bender of a pre-reader has, without benefit of special class attendance, been able to learn to read as well as her merely dull sister who has an adult type of Bender. This "miss' is of particular significance with regard to the Harriman hypothesis because reading is perhaps *the* core subject in the early grades, and of such importance that a child who is ready for it can usually fit into the rest of the program without too much difficulty.

One explanation of such an apparent contradiction to the Harriman hypothesis may be that at later ages (Nancy is now 12, remember), other factors may come into play to compensate for the inadequate visual-motor gestalt function. But what are these compensating factors? And is it necessary to assume that they become effective only at later ages?

In view of Nancy's case in which reading skill of at least third grade level was achieved in the absence of adequate Bender performance, it would seem dubious technique to use this test alone as a predictive instrument. Further study will be necessary to determine the exact extent of its value in this regard, but at present it would seem safest to use it only in conjunction with tests of general ability, and perhaps some test of reading aptitude such as Monroe's[6] which samples a variety of abilities that seem likely to contribute to the reading process.

Conclusions

1. It is pointed out that on the basis of the Harriman study of the relationship between reading skill and Bender Visual Motor Gestalt Test performance no conclusion can safely be drawn other than that both are manifestations of a general developmental process.

2. Two cases of adolescent girls are cited in which the considerably more retarded sister with an extremely im-

mature Bender performance was able to achieve a higher score on a reading test than her sister whose Bender resembled the adult type. This finding suggests the necessity of taking account of factors which may compensate for an inadequate visual-motor gestalt function.

3. It is suggested, awaiting further studies of the relationship between Bender performance and reading ability, that if it is used at all as a predictive instrument with regard to school readiness this be done only in conjunction with other tests.

Bibliography

1. Bender, L. A. *A visual motor Gestalt test and its clinical use.* New York: American Orthopsychiatric Association, 1938.
2. Gray, W. S. *Standardized oral reading paragraphs.* Bloomington, Illinois: Public School Publishing Co.
3. Harriman M. and Harriman P. L. The Bender Visual Motor Gestalt Test as a measure of school readiness. *J. clin. Psychol.*, 1950, 2, 175-177.
4. Hutt, M. A. *A mimeographed outline of instructions on the evaluative-interpretative determinants in the Bender test.* Unpublished.
5. Monroe, M. *Reading aptitude tests.* Boston, Mass.: Houghton Mifflin, 1935.

A NOTE ON THE PRESENTATION OF PERSONALITY INVENTORY SCORES

KENNETH H. IVES

New School for Social Research

With the increasing use of personality inventories, and the more frequent publication of their scores in connection with case studies, a better form for presenting these scores may be in order. Take for example the Bernreuter Personality Inventory. When used before and after treatment to show the progress made in a case, results such as given in Figure 1 are likely to be obtained. Now the author of the article in which this example appears, conveys in his text the information that the subject improved in all of the six factors between the two testings. Yet the standard form of presentation makes this improvement show as a decline in four factors, and an increase in two others. This is not conducive to clarity in the mind of the reader. It makes averaging of the results of various aspects measured a difficult operation, and it discourages efforts to correlate particular events in the interval with the change in scores.

The following rule is therefore proposed to clarify and standardize the presentation of these scores: All factors should be so defined and titled that a high score is better than a low score. This has as a corrollary that improvement in each factor brings an increased score, declines in any score represent deterioration. This rule rests partly on our customary expectation of "the more the better". A more technical basis is possible from the field of social case

Figure 1

Bernreuter Personality Inventory Scores of "Alice Taylor"; As Given in: COMBS, ARTHUR W. Follow-up of a Counseling Case Treated by the Non-directive Method. J. Clinical Psychol. 1945, Vol. 1, p. 154.

	Factor	*1943*	*1944*	*Change*
B1N	† Neurotic Tendency (high)	97	56	−41
B2S	* Self sufficiency (high)	10	25	+15
B3I	† Introversion (high) – Extraversion	95	58	−37
B4D	* Dominance (high) – Submission	4	41	+37
F1C	† Confidence in self (low)	100	68	−32
F2S	* Sociability (low)	72	45	−27

† low score desirable
* score around 50 most desirable

work. Here what is accomplished when improvement in a client occurs, often, is that the client has been helped to acquire new adaptive skills for meeting life problems.[1] The addition of skills or "assets" should logically be reflected by an addition to the inventory figure. When this rule is applied to the example previously given, the results in Figure 2 are obtained. These have the desired advantages.

FIGURE 2

THE SCORES GIVEN IN FIGURE 1 REARRANGED TO CONFORM TO THE PROCEDURE RECOMMENDED

	Factor	*1943*	*1944*	*Change*
B1N	Stability (Neurotic immunity)	3	44	41
B3I	Extraversion	5	42	37
B4D	Dominance	4	41	37
F1C	Confidence in self	0	32	32
B2S	Self sufficiency	10	25	15
F2S	Sociability	28	55	27
Averages:				
	All factors	8	40	32
	Factors with most change	3	40	37
	Factors with least change	19	40	21
	F factors only	14	44	30

Three problems arise in the application of this rule. First, where a low score was desirable in the original form, inversion of the definition must accompany inversion of the per cent figure. In the case of paired items like Dominance-Submission, this presents no problem. Factors like Neurotic Tendancy are not well handled by mere use of a negative. Either Neurotic Immunity (an interesting concept) or Stability-Instability must be used here. In a factor where the desirable score is around 50%, and the initial score in a case is far to one side of this, it must be so presented that the appropriate factor appears as weak in the first table or column. This means that in some cases the same paired factor will appear in one way, while in other cases it will appear the opposite. Thus for a domineering person the relevant factor (B4D) would be inverted. For an oversubmissive person, this same factor would not be inverted. This is a change in the basis of consistency from the factor to the case. It makes easier a direct reading of diagnostic factors and treatment needs from the scores as presented. The remaining difficulty will probably not arise often in practice. It is more obscure, and decision on it will probably have to await study of a number of cases in which it occurs. This is the case where a factor in which a score of 50% is desirable, moves from one side of this mark to the other. If it is far from the 50% mark at first, and not much on the other side at the later time, this is probably not serious. But it is not clear how to handle it if it moves from far on one side to far on the other side, or from close on one side to somewhere on the other side. These must await further investigation.

The proposed rule and its corrollaries can easily be applied to most cases where standardized personality inventories are used. It results in greater clarity of presentation, and ease in using the scores as a basis for clinical and other analysis.

REFERENCE

1. DOLLARD, JOHN and MOWRER, O. HOBART, A method of measuring tension in written documents. *J. abnorm. soc. Psychol.* 1647, 42, 4.

A NOTE ON T. A. T. ADMINISTRATION

MARY ALICE WHITE AND HANNA SCHREIBER

New York Hospital - Westchester Division

In administering the T.A.T. to psychotic and neurotic patients, we have experimented with various ways of administering the pictures in order to produce the maximum of cooperation and production. We have used the standard male or female series (1943 edition) and have been disappointed in the meagerness of the stories as well as the patient's resistance to an assigned task. We have selected certain pictures for each patient, on the basis of their Rorschach interpretation, which we hoped would clarify the dynamics in the particular case. Again, we found that adult patients of superior intelligence resent the T.A.T. as a task of some length which they have to complete, and often our purpose was defeated by obtaining only brief stories or a picture description, and we would have to resort to questioning to bring out significant material.

We believe the T.A.T. should be kept a projective technique, with a minimum of questioning, and no sense of a task attached to it. Hospitalized patients such as we see respond more easily in an informal setting where they can manipulate the situation to some extent. Therefore we have adopted the following method of administration.

1. Pictures 1 and 2 are administered first in the usual manner, asking the patient "to make up a story". We like these two standard stimuli as a basis of comparison.

2. The remainder of the cards are then handed to the patient with these instructions: "Do you want to choose those cards you'd like to tell a story about"? The patient ordinarily goes through the cards at this point, and rejections can be noted with the attending remarks. This in itself usually reveals some interesting attitudes. If the patient asks how many cards he is to choose, he is told "whatever you like". If the patient asks if he could choose none, we say "we would like you to choose the ones that interest you, but it is up to you to decide". So far, we have not had a refusal, whereas we did get refusals and partial refusals with the standard method.

We prefer this slightly different technique because it is more projective and more material is elicited. We also feel we learn more about the patient by noting which cards he chooses. We believe there is some diagnostic pattern to the choice of cards which may become apparent as our experience increases. To date we notice that male psychopaths and alcoholics choose fewer cards (as would be expected), and those often include the hobo scene (9BM), the surgeon (8BM) and son's farewell (6BM) and the "rape" scene (13MF). Rejections noted most frequently are those cards depicting passive sorrow. Usually this group makes telling sarcastic remarks about the rejected cards. It is possible that other diagnostic groups may show similar patterns.

However, the primary advantage of this modification is not diagnosis, but the increased interpretive richness of the data, and we suggest it to other examiners for their appraisal.

EDITORIAL OPINION

Among the advantages of editorial positions are the panoramic views and previews of the psychological field which are afforded to the editorial chair. Not only is an editor acquainted with the latest research trends months before they appear in print, but also one becomes aware of who is doing the work and where the most advanced centers of work are. For what such observations may be worth, your editor would like to pass along some trends which he has sensed in the course of his duties. One of the most interesting phenomena has been the waxing and waning of the prestige and influence of various university departments and training centers. In the field of clinical psychology, the most healthy leadership and activity appears to exist at some of the smaller and more recent departments rather than in the old-line eastern universities. We have been particularly impressed by the magnitude and vigor of interest in the middle and far West as revealed by number of contributions and subscriptions, particularly from student affiliates. In contrast with the vigor of clinical psychology in California, some of the Eastern departments seem moribund. We also sense that there are marked differences of interests between the older and younger generations of psychologists. Many of the younger workers show a much more intense interest in social problems, and in keeping abreast with every new development that do some of their older colleagues. There is evidence that many of the older workers are almost oblivious of the growing importance of clinical psychology as applied science and are being far outdistanced by their students. We have reached this conclusion as a result of evidence concerning who subscribes to journals and has an up-to-date knowledge of what has appeared. In professional fields of all kinds, there is a real obligation for the practitioner to keep abreast of the recent literature in order to be familiar with the latest methods of diagnosis and treatment. As an editor, we have naturally had some interest in finding out just how much is actually read and assimilated. It is very disheartening to discover that there is a lag of from five to ten years between the date of publication and the time when the material actually becomes assimilated. One of the things which we would like to see is for the APA committee on training in clinical psychology to make a frank public statement concerning its findings relating to the quality of work being done at various university departments and training centers. We venture to predict that the results might startle some of the old-line universities out of their reactionism and lethargy.

F. C. T.

In all seriousness, your editor would like to suggest the need for the establishment of a society for the advancement of semantics in professional psychology. In these days when alphabetical designations have received great popular acceptance, this new project might be called GASP (Group for the Advancement of Semantics in Psychology). Although professional psychologists usually attempt to utilize the most objective scientific methods, their modes of communication and theoretical formulations are frequently far from scientific. Probably experimental psychologists come closest to semantic validity in their operational definitions while some clinical psychologists, with their current emphasis on psychoanalytic, phenomenological and projective concepts, might be rated at the opposite extreme of semantic invalidity because of their use of symbolic terms and concepts many degrees removed from the data of basic psychological science. While attending recent psychological conventions, your editor has interested himself in appraising the semantic validity of the statements made by confreres in their scientific papers and discussions. Although viewpoints such as psychoanalysis currently enjoy great

popularity because of being espoused by many "authorities" and persons holding positions of responsibility and thereby lending respectability because of the prestige involved, many of their formulations would have to be rejected if scrutinized in the cold light of semantic and scientific objectivism. It might be wise to designate an annual Advancement of Semantics week, during which all of us would do some critical soul-searching, checking up on every statement in order to determine its validity. The conscientious utilization of semantic devices would automatically result in much more efficient and scientific modes of communication among psychologists and would eliminate much of the semantic confusion which now exists. In our editorial capacity, we have rejected many papers for semantic reasons. The authors may have understood what they were talking about but your poor editor did not. Unfortunately, these comments apply to the works of many eminent psychologists and to many of the currently popular concepts of clinical psychology. In order to qualify as true scientists, it is necessary to do more than merely join professional associations or to pass one's self off as a teacher of the subject. On purely objective grounds, many of the current efforts in clinical psychology would have to be rejected as unscientific. F. C. T.

RESEARCH SUGGESTION: It is surprising and perhaps significant that at a time when clinical psychologists are inventing new projective techniques at the rate of almost one a day, the oldest projective methods have fallen into neglect and even disrepute. To us, it has always seemed that the validation of the method of *dream interpretation* was a research problem of the highest importance. Why is it that psychoanalysts and psychiatrists have always attached the greatest significance to dream interpretation while clinical psychologists have almost ignored the method? Perhaps clinical psychologists can benefit from transfer of training in their experience with dream interpretation in applying the results to recent projective methods. It is now generally accepted that Freud's original assumption that a standard system of dream interpretation could be used to infer the latent from the manifest content has not proven valid. Apparently there is no standard vocabulary of dream symbols which has widespread validity. Does this same finding apply to Rorschach signs? It would be a very neat research problem to compare dream interpretations with Rorschach and TAT results. Three independent examiners might make blind analyses of the dreams, Rorschach and TAT results of different clinical samples of subjects. These blind analyses might be statistically analysed to discover the correlations between the findings. Another interesting problem would be to test the validity of the hypothesis that dreams tend to reflect the dominant *needs* and *presses* of the subject, i.e. the predominant emotional state whether of fear, anger or sexual deprivation. This research would not attempt to confirm the validity of specific symbolic interpretations but would simply seek to identify the basic feeling tone or emotional state coloring the dream. In our own practice, we have always placed the greatest emphasis in dream interpretation on the effort to identify dominant needs and presses.

F. C. T.

BOOK REVIEWS

THOMPSON, CLARA. *Psychoanalysis: Evolution and Development.* New York: Hermitage, 1950, pp. 252. $3.00

This short handbook presents one of the most authoritative reviews of the historical evolution of psychoanalytic theory and practice. Dr. Thompson criticizes all of the developments in psychoanalysis very impartially, discriminating valid concepts from those which have proven unfruitful. Both students and advanced workers will profit from the critical evaluations which this little book presents.

HAMILTON, KENNETH W. *Counseling the Handicapped in the Rehabilitation Process.* New York: Ronald, 1950, pp. 296. $3.50.

The author is Associate Professor of Social Administration at Ohio State University. This is an elementary textbook in the counseling of the vocationally handicapped. The nature and purpose of rehabilitation is discussed in three introductory chapters. Three chapters discuss the nature of the rehabilitation process with emphasis on the elementary mechanisms of administration, case finding, case study, vocational training and placement. Five concluding chapters deal with problems of developing community resources and evaluating the results of rehabilitation.

PEARSON, GERALD H. J. *Emotional Disorders of Children.* New York: Norton, 1949, pp. 368. $5.00.

Dr. Pearson is Director of the Philadelphia Psychoanalytic Institute and a well known psychiatrist. This is a case book of child psychiatry which includes concise discussions of the nature and treatment of the major behavior disorders of children. The viewpoint is primarily empirical with relatively few references to the recent experimental literature.

FREEMAN, FRANK S. *Theory and Practice of Psychological Testing.* New York: Holt, 1950, pp. 518. $3.50.

Dr. Freeman is Professor of Psychology at Cornell University. This is an elementary textbook suitable for introductory courses in psychological testing for all types of personnel having an interest in the field. All of the standard individual tests are described in sufficient detail for the beginning student. The viewpoint is thoroughly scientific.

KALKMAN, MARION E. *Introduction to Psychiatric Nursing.* New York: McGraw-Hill, 1950, pp. 336. $3.75.

The author is Director of Nursing and of the Illinois School of Psychiatric Nursing, Illinois Neuropsychiatric Institute, Chicago. This book will be suitable as an introductory text for nursing training schools.

BARTLEY, S. HOWARD. *Beginning Experimental Psychology.* New York: McGraw-Hill, 1950, pp. 483. $4.25.

LAWTON, GEORGE. *How to be Happy though Young.* New York: Vanguard, 1949, pp. 300. $3.00.

From 1939 to 1946, Dr. Lawton wrote a series of articles in *Scholastic Magazine* presenting answers to problems of adolescent adjustment based on unsolicited letters received by the magazine. The articles have been compiled here in a book designed to reach adolescent readers.

LEVINSON, HORACE C. *The Science of Chance.* New York: Rinehart, 1950, pp. 348. $2.00.

Those who are interested in the applications of probability theory to games of chance will welcome this book which is the best we have seen in the field.

BARTEMEIER, L. H. *et al. Third Conference on Problems of Infancy and Childhood.* New York: Josiah Macy, Jr. Foundation, 1950, pp. 156.

REIWALD, PAUL. *Society and its Criminals.* New York: International Universities Press, 1950, pp. 315. $4.50.

The author is reader in criminology in the University of Geneva. The book consists of a series of essays summarizing sociological, psychological, anthropological and psychoanalytic observations.

CHRISTENSEN, HAROLD T. *Marriage Analysis.* New York: Ronald Press, 1950, pp. 510. $4.50.

The author is Chairman of Sociology and Professor of Family Life at Purdue University. This is an elementary textbook suitable for courses in marriage and the family. While the title suggests an experimental or statistical analysis, the material consists of a general philosophical approach with rather limited reference to scientific studies in the field.

LORAND, SANDOR (Ed.) *The Yearbook of Psychoanalysis. Vol. V.* New York: International Universities Press, 1950, pp. 317. $7.50.

KATZ, DAVID. *Gestalt Psychology.* New York: Ronald, 1950, pp. 175. $3.00.

Dr. Robert Tyson has translated this reference work by the Professor of Psychology at the University of Stockholm. First appearing in 1942, it is the most authoritative critique of Gestalt psychology which has appeared in the last ten years.

LONDON, L. S. and CAPRIO, F. S. *Sexual Deviations.* Washington: Washington Institute of Medicine, 1950, pp. 702. $10.00.

This is a casebook of sexual deviations by two prominent American psychiatrists. Detailed histories are given illustrating the development background of all the major types of deviations. The viewpoint is psychoanalytic.

MAY, ROLLO. *The Meaning of Anxiety.* New York: Ronald, 1950, pp. 376. $4.50.

Dr. May is a consulting psychologist. He has reviewed the principal theories with illustrative case studies.

TERHUNE, WILLIAM B. (Ed.) *Living Wisely and Well.* New York: Dutton, 1949, pp. 95. $2.00.

This little book consists of three lectures by Douglas A. Thom, Kenneth E. Appel, and Winfred Overholser with an introduction by Dr. Terhune which were presented in 1948 by the Silver Hill Foundation. The papers discuss mental hygiene in childhood, adulthood and later life.

BAILEY, HAROLD. *The Physiology of Thought.* New York: William-Frederick Press, 1949, pp. 314. $3.75.

Dr. Bailey is a retired physician who has apparently interested himself in the psychophysiology of thought. Based on personal observations and readings, he has written a series of essays summarizing his opinions. The book is metaphysically rather than scientifically oriented. The bibliography consists of seven references dating back as far as James' *Principles of Psychology* and Ladd's *Physiological Psychology.* There is no index.

WHYTE, LANCELOT L. *The Unitary Principle in Physics and Biology.* New York: Holt, 1949, pp. 162. $3.50.

This is an important philosophical contribution expressing a conviction of the unity of nature. This unity is considered to reflect a universal form of process found in nature. Psychologists interested in a general theory of the nature of organisms will be interested in comparing this book with Lecky's Self-Consistency theory of personality unification.

FELDENKRAIS, M. *Body and Mature Behavior.* New York: International Universities Press, 1949, pp. 167. $3.75.

The subtitle of this book is "A Study of Anxiety, Sex, Gravitation and Learning." It consists of a collection of pseudoscientific essays gathering together the author's literary studies, empirical observations and conjectures on the relation between modes of posture or muscular movement and mental states.

WEITZMAN, ELLIS and MCNAMARA, WALTER, J. *Constructing Classroom Examinations.* Chicago: Science Research Associates, 1949, pp. 153. $3.00.

This is a practical handbook for teachers on all levels from elementary through university outlining the construction and use of achievement tests.

MOORE, MERRILL. *Clinical sonnets.* New York: Twayne Publishers, 1949. pp. 72. $2.50.

A leading psychiatrist presents effective pictures, in verse, of persons and their reactions to life.

READY FOR FALL CLASSES

PRINCIPLES OF PERSONALITY COUNSELING

by

FREDERICK C. THORNE, M.D.

University of Vermont

This new textbook presents a new and eclectic system of Personality Counseling, including detailed discussions and case presentations of all known methods.

PRICE: $5.00. ORDER NOW

THE MONOGRAPH SERIES

No. 1.	Billingslea: Bender Gestalt Test	$ .90
No. 2.	Carter & Bowles: Qualitative Aspects of Psychological Examining	$1.25
No. 3.	Training in Clinical Psychology	$1.50
No. 4.	Critique of Nondirective Therapy	$1.25
No. 5.	Buck: The H-T-P Manual	$2.00
No. 6.	Karpman: Objective Psychotherapy	$2.00
No. 7.	Statistics for the Clinician	$2.00

SPECIAL OFFERS

TO MEMBERS OF
THE AMERICAN PSYCHOLOGICAL
ASSOCIATION AND STUDENT AFFILIATES

1950 SUBSCRIPTION
ONLY
$4.00 [$4.50 IN CANADA]

BACK VOLUMES

COMPLETE BACK FILES
Vols. 1-5, 1945-49
ONLY $20.00

ORDER TODAY

Journal of Clinical Psychology

BRANDON, VERMONT

JOURNAL OF CLINICAL PSYCHOLOGY

EDITORIAL BOARD

FREDERICK C. THORNE, *Editor*
Medical College, University of Vermont

JERRY W. CARTER, JR.,
Wichita (Kan.) Guidance Center

SAMUEL W. HAMILTON
Essex County (N. J.) Hospital

ELAINE F. KINDER
Rockland (N. Y.) State Hospital

WILLIAM A. HUNT
Northwestern University

C. M. LOUTTIT
University of Illinois

ROBERT I. WATSON
Washington University

ELEANOR C. THORNE, *Business Manager*

INDEX

VOLUME VI

1950

JOURNAL OF CLINICAL PSYCHOLOGY
BRANDON, VERMONT

CONTENTS OF VOLUME VI

JANUARY 1950

SYMPOSIUM: STATISTICS FOR THE CLINICIAN. PAGE

Introduction: Symposium on Statistics for the Clinician. Joseph Zubin, N. Y. State Psychiatric Institute and Hospital ... 1

Statistical Problems in the Treatment of Case Histories. Alfred L. Baldwin, Fels Research Institute for the Study of Human Development, Antioch College ... 6

Experimental and Statistical Factors in the Diagnosis of Consciously Suppressed Affective Experience. Joseph F. Kubis, Fordham University ... 12

Problems of Multi-Judge Reliability. L. S. Kogan and J. McV. Hunt, Institute of Welfare Research, Community Service Society of New York ... 16

Statistical Problems Involved in Rorschach Patterning. Albert I. Rabin, Michigan State College ... 19

Statistical Methods for Multi-Score Tests. Lee J. Cronbach, Bureau of Research and Service, College of Education, University of Illinois ... 21

A Statistical Approach to Typology: The Study of Trait Universes. William Stephenson, University of Chicago ... 26

Comments on Selected Scaling Techniques with a Description of a New Type of Scale. E. F. Gardner, School of Education, Syracuse University ... 38

Intra-Individual Variability in the Study of Personality. Daniel Horn, American Cancer Society ... 43

Unitizing and Categorizing Problems in Coding Qualitative Data. Harold Guetzkow, University of Michigan ... 47

Basic Principles for Construction and Application of Discriminators. George W. Brown, The Rand Corporation, Santa Monica, California ... 58

Discussion of Symposium. John W. Tukey, Princeton University ... 61

Discussion of Symposium. John C. Flanagan, American Institute for Research, Princeton University ... 74

A Bibliography for the Student of Group Therapy. Benjamin Kotkov, Boston Regional Office, Veterans Administration ... 77

The Relation of Work Habits to Personality Adjustment. Frederick C. Thorne, University of Vermont; Lyman W. Bole, Superintendent of Schools, Springfield, Vermont; and Maurice J. O'Leary, Principal, High School, Springfield, Vermont ... 91

Editorial Opinion: Clinical Psychology in 1950 ... 101

Book Reviews ... 106

APRIL 1950

SYMPOSIUM: PSYCHOLOGY IN THE FIELD OF COMMUNITY SERVICES.

Introduction. Carlyle F. Jacobsen, State University of Iowa ... 111

The Community Services Program of the National Institute of Mental Health. Jerry W. Carter, Jr., U. S. Public Service ... 112

The Opportunities and Responsibilities of Clinical Psychology in the Field of Community Services. Keith J. Perkins, U. S Public Health Service ... 117

Experiment in Psychological Service for the Small Community. T. W. Richards, Northwestern University ... 121

Psychological Research in the Community Mental Health Service Field. Alvin Zander, University of Michigan ... 123

The Training of Psychologists for the Community Mental Health Service Field. Jean Walker MacFarlane, University of California ... 128

A Survey Evaluation of the Kinsey Report. Glenn V. Ramsey, University of Texas ... 133

On the Teaching of Clinical Psychology. William A. Hunt, Northwestern University ... 143

Research Trends in Clinical Psychology. William Schofield, University of Minnesota ... 148

Requisites for Research in Psychotherapy. Albert Ellis, The Diagnostic Center, Menlo Park, New Jersey ... 152

The Content of Mental Hygiene Literature. II. The Nature of its Directives. Robert Tyson, Hunter College ... 156

Altitude as a Reference Point in Scatter Analysis. Mordecai Whiteman, Delaware State Hospital, Farnhurst, Delaware ... 160

The Full-Range Picture Vocabulary Test: IV. Results for a White School Population. Robert B. Ammons, University of Louisville, Paul R. Arnold, University of Denver, and Robert S. Herrman, Denver Public Schools ... 164

The Bender Gestalt Test as a Projective Technique. Julian I. Kitay, Princeton University ... 170

The Bender Visual Motor Gestalt Test as a Measure of School Readiness. Mildred Harriman and Phillip Lawrence Harriman, Bucknell University ... 175

The Psychometric Localization of Brain Lesions: The Differential Effect of Frontal and Parietal Lesions on MMPI Profiles. A. Lloyd Anderson and Leo J. Hanvik, Veterans Administration Hospital, Minneapolis ... 177

Evaluating and Scoring the Minnesota Multiphasic "Cannot Say" Items. Manuel N. Brown, Veterans Administration Hospital, Vancouver, Washington ... 180
The Use of Z Scores in Equating the Wechsler-Bellevue Subtests. Irving Barnett, State University of Iowa ... 184
Wechsler-Bellevue Sub-Test Patterns: A Note of Caution. Helen E. Peixotto, Catholic University of America ... 188
The Effect of Laterality Localization of Brain Damage on Wechsler-Bellevue Indices of Deterioration. A. Lloyd Andersen, Veterans Administration Hospital, Minneapolis ... 191
The Application of the Wechsler-Bellevue Scale in the Diagnosis of Organic Brain Disorders. Brigette Gutman, Neurological Institute, New York City ... 195
A Comparative Evaluation of the Wechsler-Bellevue Mental Deterioration Index for Various Adult Groups. Lawrence S. Rogers, Veterans Administration Mental Hygiene Unit, Denver, Colorado ... 199
A Note on Allen's Index of Deterioration. Lawrence S. Rogers, Veterans Administration Mental Hygiene Unit, Denver, Colorado ... 203

Editorial Opinion ... 204
Book Reviews ... 207

JULY 1950

The Most Unpleasant Concept Test—A Graphic Projective Technique. M. R. Harrower, New York City ... 213
Research in Clinical Psychology: 1949. William Schofield, University of Minnesota ... 234
Psychiatric Opinions on Personality Factors in Psychotherapy. Ivan N. Mensh and Robert I. Watson, Washington University School of Medicine ... 237
Thematic Apperception Reactions of Crippled Children. Daniel C. Broida, Carroll E. Izard, and William M. Cruickshank, School of Education, Syracuse University ... 243
Rorschach Signs of Homosexuality in Male College Students. Leah Gold Fein, Seattle, Washington ... 248
The Prediction of Academic Success by Means of "Weighted" Harrower-Rorschach Responses. R. Travis Osborne, Wilma B. Sanders, and James E. Greene, The University of Georgia ... 253
Rorschach Characteristics of Duodenal Ulcer Patients. R. Travis Osborne and Wilma B. Sanders, University of Georgia ... 258
A Consideration of Factor Loadings on the Szondi Test. Wilson H. Guertin, Michigan State College ... 262
Alcoholic Profiles on the Minnesota Multiphasic. M. A. Brown, Mendota State Hospital ... 266
Wechsler-Bellevue Changes Following Prefrontal Lobotomy. Milton W. McCullough, Columbus State Hospital ... 270
The Prediction of Teacher-Pupil Interaction With a Projective Test. Theron Alexander, Jr., Florida State University ... 273
The Relative Difficulty of Various Tests For Sixty Feebleminded Individuals. Grace Arthur, St. Paul, Minnesota ... 276
Schizophrenia and the MMPI. Le Roy A Wauck, Chicago State Hospital and Loyola University ... 279
Changes in Goodenough Drawings Associated with Changes in Social Adjustment. Eleanore Ochs, Bellevue Psychiatric Hospital, New York ... 282
Prediction of the Response of Schizophrenic Patients to Electro-Convulsive Therapy. John S. Pearson, Veterans Administration Hospital, Minneapolis, Minnesota ... 285
Intellectual Disparities in a Normal Group and Their Relationship to Emotional Disturbance. Seymour Fisher and Helen Sunukjian, Elgin State Hospital ... 288
Animal Pictures for Obtaining Children's Projections. Robert E. Bills, University of Kentucky ... 291
A Study of the Validity of the TAT and a Set of Animal Pictures. Robert E. Bills, Charles J. Leiman, and Richard W. Thomas, University of Kentucky ... 293
An Adolescent Problem Reflected in the TAT. Leopold Bellak, Leah Levinger, and Esther Lipsky, Jewish Board of Guardians, New York City ... 295
The Nature of Rorschach Responses from Pulmonary Tuberculosis Patients. Robert W Ellis and Gladys G. Brown, Southern Methodist University ... 298
A Note on the Coefficient of Profile Similarity. Frank M. Du Mas, American Council on Education ... 300
The Use of Absolute Reliability Data in Clinical Test Interpretation. Herbert B. Weaver, University of Cincinnati ... 301

Editorial Opinion ... 303
Book Reviews ... 305

OCTOBER 1950

Intrafamilial Personality Configurations Among Children with Primary Behavior Disorders and their Parents: A Rorschach Investigation. Woodrow W. Morris and Alma L. Nichols ... 309
Objective Studies of Children's Drawings of Human Figures. I. Sex Awareness and Socio-Economic Level. Arthur Weider and Paul A. Noller, University of Louisville School of Medicine ... 315

Personality Characteristics in Hypnotically-Induced Hostility. Robert M. Counts and Ivan N. Mensh, Washington University School of Medicine 325
A Study of the Effect of Color on Productivity and Reaction Time in the Rorschach Test. Raphael J. Dubrovner, William J. Von Lackum and Hudson Jost, University of Tennessee 331
A Study of The Diagnostic Possibilities of the CVS with a Group of Diagnostic cases. Joseph D. Matarazzo, Northwestern University 337
Validation Studies of the Rorschach Test: Shyness and Gregariousness in the Normal Superior Adult. Wayne H. Holtzman, University of Texas 343
Validation Studies of the Rorschach Test: Impulsiveness in the Normal Superior Adult. Wayne H. Holtzman, University of Texas 348
Rorschach Content in Hypnosis: Chronological Age Regression. Margaret Mercer and R. W. Gibson, Veterans Administration, Coatesville, Pa. 352
A Study of Wechsler's Concept of "Normal Deterioration" in Older Mental Defectives. Gerard J. Bensberg and William Sloan, Lincoln State School and Colony 359
Judgment of Adjustment from Drawings: The Applicability of Rating Scale Methods. George W. Albee and Roy M. Hamlin University of Pittsburgh 363
The Hanfmann-Kasanin Concept Formation Test as a Measure of Rigidity in Relation to College Aptitude and Achievement. Ralph D. Norman, Charles A. Baker and Donald G. Doehring, University of New Mexico 365
Group Discussion as a Method of Patient Orientation in a Neuropsychiatric Hospital. James O. Palmer, Washington University School of Medicine 369
The Negro Patient in Psychotherapy. Ralph W. Heine, Veterans Administration, Chicago 373
An Attempted Experimental Validation of the Thematic Apperception Test. James S. Calvin, University of Kentucky, and Leo C. Ward, Veterans Administration, Atlanta 377
Szondi Results in Group Testing with College Students. David Cole and Eugene Roberts, Occidental College 381
The Test Situation and the Problem of Prediction. Seymour B. Sarason, Yale University 387
Mental Age Scores for the Wechsler Intelligence Scale for Children. William B. Grove, Phoenix Elementary Schools 393
The Wechsler-Bellevue Test and the Diagnosis of Psychopathic Personality. Milton S. Gurvitz, Hillside Hospital, Bellerose, N. Y. 397
A Study of the Differential Responses on the Vocabulary Sub-test of the Wechsler Intelligence Scale for Children. Chalmers L. Stacey and Bernard Portnoy, Syracuse University 401
Report on a Method of Predicting Success in Psychotherapy. Theodore H. Blau, Pennsylvania State College 403
Differential Scoring of the Rosenzweig Picture-Frustration Study. George Spache, Chappaqua, N. Y. 406
A Comparison of Two Methods of Administering the Digit Span Test. Richard L. Newton, University of Pittsburgh 409
A Note Regarding the Suggested Use of the Bender Visual Motor Gestalt Test as a Measure of School Readiness. Marcella Vig Baldwin, Bureau for Colored Children, Philadelphia, Pa. 412
A Note on the Presentation of Personality Inventory Scores. Kenneth H. Ives, New School for Social Research 415
A Note on T. A. T. Administration. Mary Alice White and Hanna Schreiber, New York Hospital, Westchester Division 417

Editorial Opinion 418
Book Reviews 420
Subject Index 423
Author Index 426

AUTHOR INDEX

	PAGE		PAGE		PAGE
Albee, G. W.	363	Brown, G. G.	298	Fein, L. G.	248
Alexander, T.	273	Brown, G. W.	58	Fisher, S.	288
Ammons, R. B.	164	Brown, M. A.	266	Flanagan, J.	74
Anderson, A. L.	177, 191	Brown, M. N.	180	Gardner, E. F.	38
Arnold, P. R.	164	Calvin, J. S.	377	Gibson, R. W.	352
Arthur, G.	276	Carter, J. W.	112	Greene, J. E.	253
Baker, C. A.	365	Cole, D.	381	Groves, W. B.	393
Baldwin, A. L.	6	Counts, R. M.	325	Guertin, W. H	262
Baldwin, M. V.	412	Cronbach, L. J	21	Guetzkow, H.	47
Barnett, I.	184	Cruickshank, W. M.	243	Gutman B.	195
Bellak, L.	295	Doehring, D. G.	365	Hamlin, R. M.	363
Bensberg, G. J.	359	Dubrovner, R. J.	331	Harriman, M.	175
Blau, T. H.	403	Du Mas, F. M.	300	Harriman, P. L.	175
Bole, L. S.	91	Ellis, A.	152	Harrower, M. R.	213
Broida, D. C.	243	Elis, R. W.	298	Heine, R. W.	373

(Continued on next page)

AUTHOR INDEX (*Concluded*)

	PAGE
Hermann, R. S.	164
Holtzman, W. H.	343, 348
Horn, D.	43
Hunt, J. McV.	16
Hunt, W. A.	143
Ives, K. H.	415
Izard, C. E.	243
Jacobsen, C. F.	111
Jost, H.	331
Kitay, J. I.	170
Kogan, L. S.	16
Kotkov, B.	77
Kubis, J. F.	12
Leiman, C. J.	293
Levinger, L.	295
Lipsky, E.	295
MacFarlane, J. W.	128
Matarazzo, J. D.	337
McCullough, M. W.	270
Mensh, I. N.	237, 325
Mercer, M.	352
Morris, W. W.	309
Newton, R. L.	409
Nichols, A. L.	309
Noller, P. A.	315
Norman, R. D.	365
Ochs, E.	282
O'Leary, M. J.	91
Osborne, R. T.	253, 258
Palmer, J. O.	369
Pearson, J. S.	285
Perkins, K. J.	117
Portnoy, B.	401
Rabin, A. I.	19
Ramsey, G. V.	133
Richards, T. W.	121
Roberts, E.	381
Rogers, L. S.	199, 203
Sanders, W. B.	258
Sarason, S. B.	387
Schofield, W.	148, 234
Schreiber, H.	417
Sloan, W.	359
Spache, G.	406
Stacey, C. L.	401
Stephenson, W.	26
Sunukjian, H.	288
Thomas, R. W.	293
Thorne, F. C.	91
Tukey, J. W.	61
Tyson, R.	156
Von Lackum, W. J.	331
Ward, L. C.	377
Watson, R. I.	237
Wauck, L.	279
Weaver, H. B.	301
Weider, A.	315
White, M. A.	417
Whiteman, M.	160
Zander, A.	123
Zubin, J.	1

Editorial Opinion 101, 204, 303, 418
Book Reviews 106, 207, 305, 420

PHILOSOPHICAL LIBRARY
PUBLICATIONS

FREUD: Dictionary of Psychoanalysis

Edited by Nandor Fodor and Frank Gaynor

With a Preface by Theodor Reik

Here for the first time, fulfilling a need urgently felt by the specialist and student alike, is a complete glossary of all the basic terms in psychoanalysis.

The concise definitions, each taken word for word from the psychological writings of Dr. Freud, cover all the basic terms and concepts in this new science of the human mind. These most informative definitions, as Freud himself formulated them, are incorporated in this handy one-volume dictionary.

A genuine storehouse of information for the student, this volume is at the same time an extremely convenient and useful reference work for the psychologist, psychiatrist, and psychoanalyst.

$3.75

"May well find its place on library shelves among other excellent recent volumes."

—V. C. Branham, M.D.
Chief, Outpatient Section
Psychiatry and Neurology Division
Veterans Administration, Wash. D. C.

MODERN ABNORMAL PSYCHOLOGY

Edited by William H. Mikesell

23 authorities (including Lyon, Zilboorg, Fleming, Brill, Adams) treat all aspects exhaustively in terms understandable to both students, scholars, and practitioners. Causes of mental diseases are clearly explained, diagnostic procedures fully described and treatment techniques impartially discussed: Freudian, Neo-Freudian, and Non-Analytical Therapy. In addition, stress is put on the role of psychosomatic medicine.

$10.00

NEW DICTIONARY OF PSYCHOLOGY 2nd EDITION

By Philip Lawrence Harriman

No effort has been spared to include all basic terms germane to modern psychology, psychoanalysis, psychiatry and allied subjects. In this rapidly expanding field, with its constantly growing vocabulary, the new concepts require a clearly-defined terminology based upon up-to date research and contemporary literature.

$5.00

Expedite shipment by prepayment

Special student bulk rate on 10 or more

PHILOSOPHICAL LIBRARY

15 East 40th Street, Desk 190 New York 16, N. Y.

PROVIDES PHYSICIANS WITH FACILITIES FOR

The Education and Treatment of Children With Academic or Emotional Difficulties

Applications for fall entrance now being received. A representative of Devereux may soon be in your vicinity and would be glad to confer with you or with any parent you recommend Write the Registrar.

Twelve separate home-school units in Pennsylvania and the Devereux Ranch School in California offer a controlled environment and modern training shaped to the needs of each child.

Devereux is unique in that it combines complete school facilities for children of kindergarten through high school age with complete medical, psychiatric, and psychological services.

The professional staff cooperates with the referring physician in his plans for the child's training and care.

HELEN T. DEVEREUX, *Director*

Resident Medical Staff

RUTH E. DUFFY, M.D.
ROBERT T. DEVEREUX, M.D.
LESLIE R. ANGUS, M.D.
ALEXANDER AARONS, M.D.

ROBERT G. FERGUSON, A.M., *Educational Psychologist*
MICHAEL B. DUNN, A.M., *Clinical Psychologist*

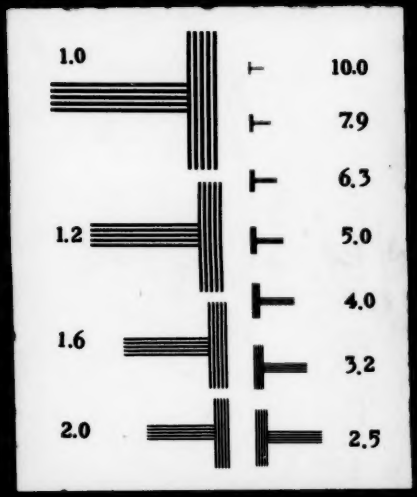
EDGAR A. DOLL, PH.D., *Coordinator of Research*

For Information, address the Registrar
JOHN M. BARCLAY, Deven, Pennsylvania

Under the Devereux Foundation, an organization not for profit

Santa Barbara, California **Devon, Pennsylvania**

RESOLUTION CHART



100 MILLIMETERS

INSTRUCTIONS Resolution is expressed in terms of the lines per millimeter recorded by a particular film under specified conditions. Numerals in chart indicate the number of lines per millimeter in adjacent "T-shaped" groupings.

In microfilming, it is necessary to determine the reduction ratio and multiply the number of lines in the chart by this value to find the number of lines recorded by the film. As an aid in determining the reduction ratio, the line above is 100 millimeters in length. Measuring this line in the film image and dividing the length into 100 gives the reduction ratio. Example: the line is 20 mm. long in the film image, and $100/20 = 5$.

Examine "T-shaped" line groupings in the film with microscope, and note the number adjacent to finest lines recorded sharply and distinctly. Multiply this number by the reduction factor to obtain resolving power in lines per millimeter. Example: 7.9 group of lines is clearly recorded while lines in the 10.0 group are not distinctly separated. Reduction ratio is 5, and $7.9 \times 5 = 39.5$ lines per millimeter recorded satisfactorily. $10.0 \times 5 = 50$ lines per millimeter which are not recorded satisfactorily. Under the particular conditions, maximum resolution is between 39.5 and 50 lines per millimeter.

Resolution, as measured on the film, is a test of the entire photographic system, including lens, exposure, processing, and other factors. These rarely utilize maximum resolution of the film. Vibrations during exposure, lack of critical focus, and exposures yielding very dense negatives are to be avoided.

THIS PUBLICATION IS REPRO-
DUCED BY AGREEMENT WITH THE
COPYRIGHT OWNER. EXTENSIVE
DUPLICATION OR RESALE WITH-
OUT PERMISSION IS PROHIBITED.